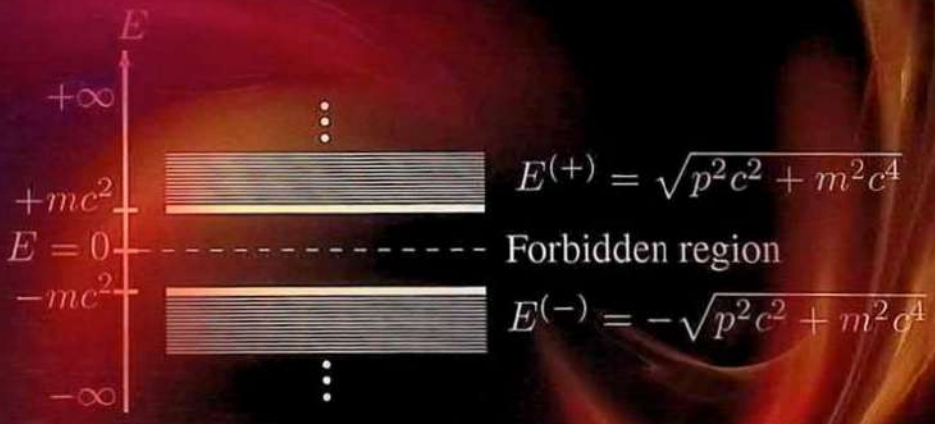
THIRD EDITION

QUANTUM MECHANICS

CONCEPTS AND APPLICATIONS

NOUREDINE ZETTILI



WILEY

Quantum Mechanics

Concepts and Applications

Third Edition

Nouredine Zettili
Jacksonville State University

July 1, 2022

WILEY

This 3rd edition first published 2022
© 2022 John Wiley & Sons Ltd

Edition History
2nd edition, published 2009 by John Wiley & Sons, Ltd
1st edition, published 2001 by John Wiley & Sons, Ltd

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at http://www.wiley.com/go/permissions.

The right of Nouredine Zettili to be identified as the author of this work has been asserted in accordance with law.

Registered Offices
John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Office
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty
While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data

Name: Zettili, Nouredine, author.
Title: Quantum mechanics : concepts and applications – 3rd ed. / Nouredine Zettili, Jacksonville State University, USA.

Description: Third edition. / Chichester, UK: Wiley, 2022 / Includes bibliographical references and index.
Identifiers: ISBN 9781118307892 (pbk.)
Subjects: Quantum theory.
Classification: QC174.12.Z47 2022 / 530.12–dc22

Cover Design: Wiley
Cover image: © ALFRED PASIEKA/Getty Images

Produced from LaTeX files supplied by the author
Printed and bound by CPI Group (UK) Ltd, Croydon, CR0 4YY

C9781118307892_030822

Contents

1 **Origins of Quantum Physics** 1
1.1 Historical Note: From Classical to Modern Physics . . . 1
1.1.1 1859–1900: Failure of Classical Physics at the Microscopic Scale . . . 2
1.1.2 1900–1932: Advent of Quantum Mechanics – New Era in Physics . . . 4
1.2 Particle Aspect of Radiation . . . 12
1.2.1 Blackbody Radiation . . . 13
1.2.2 Photoelectric Effect . . . 19
1.2.3 Compton Effect . . . 23
1.2.4 Pair Production . . . 25
1.3 Wave Aspect of Particles . . . 27
1.3.1 de Broglie's Hypothesis: Matter Waves . . . 27
1.3.2 Experimental Confirmation of de Broglie's Hypothesis . . . 28
1.3.3 Matter Waves for Macroscopic Objects . . . 30
1.4 Particles versus Waves . . . 31
1.4.1 Classical View of Particles and Waves . . . 31
1.4.2 Quantum View of Particles and Waves . . . 33
1.4.3 Wave–Particle Duality: Complementarity . . . 35
1.4.4 Principle of Linear Superposition . . . 37
1.5 Indeterministic Nature of the Microphysical World . . . 38
1.5.1 Heisenberg's Uncertainty Principle . . . 38
1.5.2 Probabilistic Interpretation . . . 40
1.6 Bohr's Atomic Model, Atomic Transitions and Spectroscopy . . . 40
1.6.1 Rutherford Planetary Model of the Atom . . . 41
1.6.2 Bohr Model of the Hydrogen Atom . . . 41
1.7 Quantization Rules . . . 48
1.8 Wave Packets . . . 50
1.8.1 Localized Wave Packets . . . 51
1.8.2 Wave Packets and the Uncertainty Relations . . . 54
1.8.3 Motion of Wave Packets . . . 55
1.9 Concluding Remarks . . . 66
1.10 Solved Problems . . . 67
1.11 Exercises . . . 84

2 Mathematical Tools of Quantum Mechanics 91
2.1 Introduction . . . 91
2.2 The Hilbert Space and Wave Functions . . . 91
2.2.1 The Linear Vector Space . . . 91
2.2.2 The Hilbert Space . . . 92
2.2.3 Dimension and Basis of a Vector Space . . . 93
2.2.4 Square-Integrable Functions: Wave Functions . . . 96
2.3 Dirac Notation . . . 97
2.4 Operators . . . 104
2.4.1 General Definitions . . . 104
2.4.2 Hermitian Adjoint . . . 105
2.4.3 Projection Operators . . . 107
2.4.4 Commutator Algebra . . . 108
2.4.5 Uncertainty Relation between Two Operators . . . 110
2.4.6 Functions of Operators . . . 112
2.4.7 Inverse of an Operator and Unitary Operators . . . 113
2.4.8 Antilinear and Antiunitary Operators, Wigner Theorem . . . 114
2.4.9 Eigenvalues and Eigenvectors of an Operator . . . 118
2.4.10 Infinitesimal and Finite Unitary Transformations . . . 120
2.5 Representation in Discrete Bases . . . 123
2.5.1 Matrix Representation of Kets, Bras, and Operators . . . 123
2.5.2 Change of Bases and Unitary Transformations . . . 134
2.5.3 Matrix Representation of the Eigenvalue Problem . . . 137
2.6 Representation in Continuous Bases . . . 141
2.6.1 General Treatment . . . 141
2.6.2 Position Representation . . . 142
2.6.3 Momentum Representation . . . 143
2.6.4 Connecting the Position and Momentum Representations . . . 144
2.6.5 Parity Operator . . . 148
2.7 Matrix and Wave Mechanics . . . 150
2.7.1 Matrix Mechanics . . . 151
2.7.2 Wave Mechanics . . . 151
2.8 Concluding Remarks . . . 152
2.9 Solved Problems . . . 153
2.10 Exercises . . . 176

3 Postulates of Quantum Mechanics 187
3.1 Introduction . . . 187
3.2 The Basic Postulates of Quantum Mechanics . . . 187
3.3 The State of a System . . . 190
3.3.1 Probability Density . . . 190
3.3.2 The Superposition Principle . . . 191
3.4 Observables and Operators . . . 193
3.5 Measurement in Quantum Mechanics . . . 195
3.5.1 How Measurements Disturb Systems . . . 195
3.5.2 Expectation Values . . . 196
3.5.3 Complete Sets of Commuting Operators (CSCO) . . . 198

3.5.4 Measurement and the Uncertainty Relations . . . 200
3.6 Time Evolution of the System's State . . . 201
3.6.1 Time Evolution Operator . . . 201
3.6.2 Stationary States: Time-Independent Potentials . . . 202
3.6.3 Schrödinger Equation and Wave Packets . . . 203
3.6.4 The Conservation of Probability . . . 204
3.6.5 Time Evolution of Expectation Values . . . 206
3.7 Symmetries and Conservation Laws . . . 207
3.7.1 Infinitesimal Unitary Transformations . . . 207
3.7.2 Finite Unitary Transformations . . . 208
3.7.3 Symmetries and Conservation Laws . . . 209
3.8 Connecting Quantum to Classical Mechanics . . . 211
3.8.1 Poisson Brackets and Commutators . . . 211
3.8.2 The Ehrenfest Theorem . . . 213
3.8.3 Quantum Mechanics and Classical Mechanics . . . 214
3.9 Solved Problems . . . 216
3.10 Exercises . . . 234

4 One-Dimensional Problems 241
4.1 Introduction . . . 241
4.2 Properties of One-Dimensional Motion . . . 242
4.2.1 Discrete Spectrum (Bound States) . . . 242
4.2.2 Continuous Spectrum (Unbound States) . . . 243
4.2.3 Mixed Spectrum . . . 243
4.2.4 Symmetric Potentials and Parity . . . 244
4.3 The Free Particle: Continuous States . . . 244
4.4 The Potential Step . . . 246
4.5 The Potential Barrier and Well . . . 250
4.5.1 The Case $E > V_0$ . . . 250
4.5.2 The Case $E < V_0$: Tunneling . . . 254
4.5.3 The Tunneling Effect . . . 257
4.6 The Infinite Square Well Potential . . . 258
4.6.1 The Asymmetric Square Well . . . 258
4.6.2 The Symmetric Potential Well . . . 261
4.7 The Finite Square Well Potential . . . 262
4.7.1 The Scattering Solutions ($E > V_0$) . . . 262
4.7.2 The Bound State Solutions ($0 < E < V_0$) . . . 262
4.8 The Harmonic Oscillator . . . 266
4.8.1 Energy Eigenvalues . . . 269
4.8.2 Energy Eigenstates . . . 271
4.8.3 Energy Eigenstates in Position Space . . . 272
4.8.4 The Matrix Representation of Various Operators . . . 275
4.8.5 Expectation Values of Various Operators . . . 276
4.9 Numerical Solution of the Schrödinger Equation . . . 277
4.9.1 Numerical Procedure . . . 277
4.9.2 Algorithm . . . 279
4.10 Solved Problems . . . 281

4.11 Exercises 305

5 Angular Momentum **313**
5.1 Introduction 313
5.2 Orbital Angular Momentum 313
5.3 General Formalism of Angular Momentum 315
5.4 Matrix Representation of Angular Momentum 320
5.5 Geometrical Representation of Angular Momentum 323
5.6 Spin Angular Momentum 325
5.6.1 Experimental Evidence of the Spin 325
5.6.2 General Theory of Spin 329
5.6.3 Spin 1/2 and the Pauli Matrices 330
5.6.4 Schrödinger Equation with Spin: Pauli Hamiltonian 333
5.7 Eigenfunctions of the Orbital Angular Momentum 335
5.7.1 Eigenfunctions and Eigenvalues of $\hat{L}_z$ 336
5.7.2 Eigenfunctions of $\hat{\vec{L}}^2$ 336
5.7.3 Properties of the Spherical Harmonics 340
5.8 Solved Problems 344
5.9 Exercises 359

6 Three-Dimensional Problems **367**
6.1 Introduction 367
6.2 3D Problems in Cartesian Coordinates 367
6.2.1 General Treatment: Separation of Variables 367
6.2.2 The Free Particle 369
6.2.3 The Box Potential 370
6.2.4 The Harmonic Oscillator 372
6.3 3D Problems in Spherical Coordinates 374
6.3.1 Central Potential: General Treatment 374
6.3.2 The Free Particle in Spherical Coordinates 378
6.3.3 The Spherical Square Well Potential 380
6.3.4 The Isotropic Harmonic Oscillator 382
6.3.5 The Hydrogen Atom 385
6.3.6 Effect of Magnetic Fields on Central Potentials 400
6.4 Pauli Equation for a Spin 1/2 Particle in a Magnetic Field 403
6.4.1 Pauli equation for a free Spin 1/2 Particle 404
6.4.2 Pauli equation for a Spin 1/2 Particle in an Electromagnetic Field 405
6.5 Concluding Remarks 408
6.6 Solved Problems 410
6.7 Exercises 427

7 Rotations and Addition of Angular Momenta **433**
7.1 Rotations in Classical Physics 433
7.1.1 Active Rotations 433
7.1.2 Passive Rotations 434
7.1.3 Properties of Rotation Matrices 436
7.2 Rotations in Quantum Mechanics 437

7.2.1 Infinitesimal Rotations . . . 438
7.2.2 Finite Rotations . . . 439
7.2.3 Properties of the Rotation Operator . . . 440
7.2.4 Euler Rotations . . . 442
7.2.5 Representation of the Rotation Operator . . . 443
7.2.6 Rotation Matrices and the Spherical Harmonics . . . 445
7.3 Addition of Angular Momenta . . . 448
7.3.1 Addition of Two Angular Momenta: General Formalism . . . 448
7.3.2 Calculation of the Clebsch–Gordan Coefficients . . . 454
7.3.3 Coupling of Orbital and Spin Angular Momenta . . . 460
7.3.4 Addition of More Than Two Angular Momenta . . . 464
7.3.5 Rotation Matrices for Coupling Two Angular Momenta . . . 465
7.3.6 Isospin . . . 468
7.4 Rotations of Scalar, Vector, and Tensor Operators . . . 471
7.4.1 Transformation of Scalar Operators Under Rotations . . . 471
7.4.2 Transformation of Vector Operators Under Rotations . . . 472
7.4.3 Transformation of Tensor Operators Under Rotations . . . 473
7.4.4 Wigner–Eckart Theorem for Spherical Tensor Operators . . . 476
7.5 Solved Problems . . . 479
7.6 Exercises . . . 496

8 Identical Particles 503
8.1 Many-Particle Systems . . . 503
8.1.1 Schrödinger Equation . . . 503
8.1.2 Interchange Symmetry . . . 505
8.1.3 Systems of Distinguishable Noninteracting Particles . . . 506
8.2 Systems of Identical Particles . . . 508
8.2.1 Identical Particles in Classical and Quantum Mechanics . . . 508
8.2.2 Exchange Degeneracy . . . 510
8.2.3 Symmetrization Postulate . . . 511
8.2.4 Constructing Symmetric and Antisymmetric Functions . . . 512
8.2.5 Systems of Identical Noninteracting Particles . . . 513
8.3 The Pauli Exclusion Principle . . . 516
8.4 The Exclusion Principle and the Periodic Table . . . 517
8.5 Solved Problems . . . 524
8.6 Exercises . . . 533

9 Approximation Methods for Stationary States 537
9.1 Introduction . . . 537
9.2 Time-Independent Perturbation Theory . . . 538
9.2.1 Nondegenerate Perturbation Theory . . . 538
9.2.2 Degenerate Perturbation Theory . . . 544
9.2.3 Fine Structure and the Anomalous Zeeman Effect . . . 548
9.3 The Variational Method . . . 558
9.4 The Wentzel–Kramers–Brillouin Method . . . 566
9.4.1 General Formalism . . . 566
9.4.2 Bound States for Potential Wells with No Rigid Walls . . . 569

9.4.3 Bound States for Potential Wells with One Rigid Wall 575
9.4.4 Bound States for Potential Wells with Two Rigid Walls 577
9.4.5 Tunneling through a Potential Barrier 579
9.5 Concluding Remarks 582
9.6 Solved Problems 582
9.7 Exercises 614

10 Time-Dependent Perturbation Theory **623**
10.1 Introduction 623
10.2 The Pictures of Quantum Mechanics 623
10.2.1 The Schrödinger Picture 624
10.2.2 The Heisenberg Picture 624
10.2.3 The Interaction Picture 625
10.3 Time-Dependent Perturbation Theory 627
10.3.1 Transition Probability 629
10.3.2 Transition Probability for a Constant Perturbation 630
10.3.3 Transition Probability for a Harmonic Perturbation 631
10.4 Adiabatic and Sudden Approximations 635
10.4.1 Adiabatic Approximation 635
10.4.2 Sudden Approximation 636
10.5 Interaction of Atoms with Radiation 639
10.5.1 Classical Treatment of the Incident Radiation 640
10.5.2 Quantization of the Electromagnetic Field 641
10.5.3 Transition Rates for Absorption and Emission of Radiation 644
10.5.4 Transition Rates within the Dipole Approximation 645
10.5.5 The Electric Dipole Selection Rules 646
10.5.6 Spontaneous Emission 647
10.6 Solved Problems 650
10.7 Exercises 667

11 Scattering Theory **671**
11.1 Scattering and Cross Section 671
11.1.1 Connecting the Angles in the Lab and CM frames 672
11.1.2 Connecting the Lab and CM Cross Sections 674
11.2 Scattering Amplitude of Spinless Particles 675
11.2.1 Scattering Amplitude and Differential Cross Section 678
11.2.2 Scattering Amplitude 678
11.3 The Born Approximation 682
11.3.1 The First Born Approximation 682
11.3.2 Validity of the First Born Approximation 683
11.4 Partial Wave Analysis 685
11.4.1 Partial Wave Analysis for Elastic Scattering 685
11.4.2 Partial Wave Analysis for Inelastic Scattering 689
11.5 Scattering of Identical Particles 691
11.6 Solved Problems 693
11.7 Exercises 704

12 Relativistic Quantum Mechanics **707**
12.1 Introduction . . . 707
12.2 Four Vectors and Minkowski Metric . . . 709
12.2.1 Euclidean Metric . . . 709
12.2.2 Minkowski Spacetime and the Lorentz Transformation . . . 710
12.2.3 Four-Vectors . . . 711
12.2.4 Minkowski Metric . . . 712
12.3 Klein–Gordon Equation . . . 715
12.3.1 Historical note about the Genesis of the Klein–Gordon Equation . . . 715
12.3.2 Klein–Gordon Equation for a Free particle and its Covariance . . . 717
12.3.3 Nonrelativistic Limit of the Klein–Gordon Equation . . . 719
12.3.4 Conceptual Problems of the Klein–Gordon Equation . . . 721
12.3.5 Klein Paradox for Spin-Zero Particles . . . 725
12.3.6 Klein–Gordon Equation in an Electromagnetic Field . . . 731
12.3.7 Non-Relativistic Limit of the KGE in an Electromagnetic Field . . . 732
12.4 Dirac Equation . . . 734
12.4.1 Dirac Equation for a Free Particle . . . 735
12.4.2 Representation of the Dirac Matrices . . . 737
12.4.3 Dirac Matrices and their Properties . . . 739
12.4.4 Dirac, Weyl, and Majorana Representations of the γ Matrices . . . 746
12.4.5 Dirac Equation in Covariant Form . . . 748
12.4.6 Lorentz Covariance of the Dirac Equation . . . 749
12.4.7 Continuity Equation and Current Conservation . . . 753
12.4.8 Solutions of the Dirac Equation for a Free Particle at Rest . . . 756
12.4.9 Solutions of the Dirac Equation for a Free Particle in Motion . . . 760
12.4.10 Negative-Energies Interpretation: Dirac's Hole Theory . . . 763
12.4.11 Negative-Energies Interpretation: Feynman-Stückelberg Method . . . 766
12.4.12 Negative-Energies Interpretation: Klein Paradox for Spin-1/2 Particles . 773
12.4.13 Spin-Half Solutions of the Dirac Equation . . . 777
12.4.14 Helicity . . . 782
12.4.15 Dirac Equation for Massless Particles: Weyl Fermions . . . 784
12.4.16 Weyl Spinors and Parity Violation . . . 786
12.4.17 Weyl Fermions and Neutrinos . . . 786
12.4.18 Chirality . . . 789
12.4.19 Dirac Equation in an Electromagnetic Field . . . 793
12.4.20 Non-relativistic limit of the Dirac Equation . . . 794
12.4.21 Discrete Symmetries of the Dirac Equation . . . 800
12.4.22 Charge Conjugation and Covariance of the Dirac Equation . . . 800
12.4.23 Parity and Covariance of the Dirac Equation . . . 805
12.4.24 Time-Reversal and Covariance of the Dirac Equation . . . 810
12.4.25 The *CPT* Theorem . . . 815
12.5 Successes and Limitations of Relativistic Quantum Mechanics . . . 817
12.6 Beyond Relativistic Quantum Mechanics . . . 819
12.7 Solved Problems . . . 819
12.8 Exercises . . . 889

13 Beyond Relativistic Quantum Mechanics 897
13.1 Introduction . . . 897
13.2 Advantages of Lagrangians in Field Theories . . . 898
13.3 Lagrangians in Classical Field Theory . . . 899
13.4 Lagrangian Derivation of the Maxwell Equations . . . 903
13.4.1 Maxwell Equations in Differential Form . . . 903
13.4.2 Maxwell Equations in Covariant Form . . . 905
13.4.3 Lagrangian Derivation of the Maxwell Equations . . . 909
13.5 Lagrangian Derivation of the Klein-Gordon Equation . . . 911
13.6 Lagrangian Derivation of the Dirac Equation . . . 915
13.7 Beyond Classical Field Theory: Quantum Field Theory . . . 919
13.8 The Lagrangian of Quantum Electrodynamics . . . 920
13.9 Concluding Thoughts . . . 922
13.10Solved Problems . . . 923
13.11Exercises . . . 934

A The Delta Function 937
A.1 One-Dimensional Delta Function . . . 937
A.1.1 Various Definitions of the Delta Function . . . 937
A.1.2 Properties of the Delta Function . . . 938
A.1.3 Derivative of the Delta Function . . . 939
A.2 Three-Dimensional Delta Function . . . 940

B Angular Momentum in Spherical Coordinates 941
B.1 Derivation of Some General Relations . . . 941
B.2 Gradient and Laplacian in Spherical Coordinates . . . 942
B.3 Angular Momentum in Spherical Coordinates . . . 943

C C++ Code for Solving the Schrödinger Equation 945
C.1 Introduction . . . 945
C.2 C++ Code for Solving the Schrödinger Equation . . . 945

D Index Notation for 4-Vectors 951
D.1 Index Notation for 3 and 4-Vectors . . . 951
D.1.1 Introductory Reminder . . . 951
D.1.2 Index Notation for 3-Vectors . . . 953
D.1.3 Inner and Cross Products of 3-Vectors in the Index Notation . . . 954
D.1.4 Index Notation for 4-Vectors . . . 956
D.2 The Minkowski Metric . . . 957

E The Relativistic Notation and Four Vectors 961
E.1 Introduction . . . 961
E.2 Lorentz transformation . . . 962
E.3 Minkowski Spacetime . . . 964
E.4 The 4-Vector Notation . . . 964
E.5 Contravariant and Covariant notations . . . 968
E.6 Lorentz transformation in the contravariant-covariant form . . . 974

E.7 Invariance under the Lorentz transformation 977

F Lagrangian Formulation of Classical Mechanics **979**
F.1 Lagrangian Formulation of Classical Mechanics 979
F.2 Hamiltonian Formulation of Classical Mechanics 984
F.3 Canonical Quantization: from Classical to Quantum Mechanics 989

Index **993**

Preface

Preface to the Third Edition

The second edition of this book has appeared about 13 years ago. During this time, I have received consistent feedback from worldwide, courteous users of the book; literally, users from every continent. The most persistent request from my readers was to add a chapter about relativistic quantum mechanics. So the main aim behind this new edition is to address this need. As a result, I have added two new chapters – Chapters 12 and 13 – and three appendices – Appendices D, E, F – to support them. Following the same approach of the second edition, these two new chapters contain rich collections of examples, solved problems, and exercises aimed at enabling the users to master the technical aspects of relativistic quantum mechanics and classical field theory. In addition to these newly added chapters, I have implemented a number of substantive modifications to some chapters so as to update them and to refine and streamline their contents.

In a nutshell, the new materials added to the 3rd edition can be summarized as follows:

1. Chapter 12: *Relativistic Quantum Mechanics*, dealing primarily with the various aspects of the Klein-Gordon and Dirac equations.

2. Chapter 13: *Beyond Relativistic Quantum Mechanics*, outlining some of the most essential elements of classical field theory with an emphasis on the Lagrangian derivations of the Klein-Gordon and Dirac equations.

3. Appendix D: *Index Notation for 4-Vectors*, covering some much needed notations that make the formalisms of chapters 12 and 13 as well as Appendix E more transparent.

4. Appendix E: *The Relativistic Notation and 4-Vectors*, treating the most needed technical aspects to render the contents of chapters 12 and 13 easy to follow.

5. Appendix F: *Lagrangian and Hamiltonian Formulations of Classical Mechanics*, covering the technical as well as conceptual contents that are necessary to follow Chapter 13.

6. A complete overhaul of Section 1.1: *Historical Note: From Classical to Modern Physics*, offering an in-depth coverage of the most important ideas that led to the birth of quantum mechanics, which, eventually, paved the way for the creation and development of the various fields of modern physics.

7. Several additions to Chapter 2, most notably, Subsection 2.4.8: *Antilinear and Antiunitary Operators, Wigner Theorem*, covering topics that are needed in Chapter 12.

8. A new section: Section 6.4: *Pauli Equation for a Spin 1/2 Particle in a Magnetic Field.*

9. Major overhaul of Sections 7.1 and 7.2.

10. I have added the most important, historical references to the topics covered in chapters 1, 12, and 13.

11. I have restructured the formulas and equations throughout the entire text to make them compliant with the SI units; so the 3rd editions uses SI, not cgs, units.

With these new additions, especially chapters 12 and 13, the 3rd edition offers sufficient contents for a two-semester undergraduate course as well as for a one or two graduate courses in quantum mechanics.

Acknowledgments

Two seasoned, world-class physicists – Prof. Ismail Zahed (Stony Brrok University, New York) and Prof. Salah Nasri (UAE University, El-Ain) – have courteously accepted and carried out the punishing task of reviewing Chapters 12 and 13 along with appendices D, E, F. I want to profusely thank both of them for having generously taken the time and care to comb through these chapters and appendices and checked their accuracy. I will be immensely indebted to both of them. Special thanks are also due to Prof. Robert Jafee, my former professor during my graduate studies at MIT, for having read chapters 12 and 13 and for having given me several pertinent suggestions.

Over the course of the last 12 years, I have received useful feedback from worldwide users of the book— professors, researchers, students, and retired scientists. To each one of them, I offer my most special thanks. Although the list is too long to mention, I want to thank in particular the following courteous users who have provided me with praiseworthy input: Prof. Christian Flindt (Aalto University, Finland), Prof. Mathew Leifer (Chapman University, USA), Prof. Ricardo Kagimura (Federal University of Uberlandia, Brasil), Dr. Jesni Shamsul (Malaysia), Prof. Christopher Crawford (University of Kentucky, USA), Prof. Siddhesh Padwal (India), Dr.Mohammad Amid (Iran), Dr. Ahmad Humsi (Denmark), Dr. Muhammad Naeem (India), Prof. Murat Hazer Uygunol (Edge University, Turkey), Dr. Håkon Hallingstad (Norway), Prof. Wa'el Salah (Jordan), Prof. N. Shaji (Maharajas's College, India), Prof. Mubarak Aydh (Saudi Arabia), Prof. Maien Binjonaid (Saudi Arabia), Prof. Camilo Estrada Guerra (National University of Colombia, Columbia), Dr. Kamal Joshi (Iowa State University, USA), Prof. Suhail A. Siddiqui (AMU, India), Prof. Atul Choudhary (India), Dr. Mevlut Bulut (University of Alabama, Birmingham, USA), Dr. Naman Agarwal (Germany), and Prof. Alexander O. Govorov (Ohio University, Athens, USA).

Finally, I want to thank Ms. Jenny Cosshamm, my Publisher and the Associate Editorial Director, for having patiently waited 12 years for the 3rd edition to see the daylight! Her unwavering support, leadership, and skillful management of the various phases of this project are highly appreciated. Special thanks are also due to Dr. Martin Preuss, my new Publisher and the Physics Associate Publisher, and Ms. Amy Odum, Managing Editor, for their support in concluding this project.

N. Zettili
Jacksonville State University
May 2022

Preface to the Second Edition

It has been eight years now since the appearance of the first edition of this book in 2001. During this time, many courteous users—professors who have been adopting the book, researchers, and students—have taken the time and care to provide me with valuable feedback about the book. In preparing the second edition, I have taken into consideration the generous feedback I have received from these users. To them, and from the very outset, I want to express my deep sense of gratitude and appreciation.

The underlying focus of the book has remained the same: to provide a well-structured and self-contained, yet concise, text that is backed by a rich collection of fully solved examples and problems illustrating various aspects of nonrelativistic quantum mechanics. The book is intended to achieve a double aim: on the one hand, to provide instructors with a pedagogically suitable teaching tool and, on the other, to help students not only master the underpinnings of the theory but also become effective practitioners of quantum mechanics.

Although the overall structure and contents of the book have remained the same upon the insistence of numerous users, I have carried out a number of streamlining, surgical type changes in the second edition. These changes were aimed at fixing the weaknesses (such as typos) detected in the first edition while reinforcing and improving on its strengths. I have introduced a number of sections, new examples and problems, and new material; these are spread throughout the text. Additionally, I have operated substantive revisions of the exercises at the end of the chapters; I have added a number of new exercises, jettisoned some, and streamlined the rest. I may underscore the fact that the collection of end-of-chapter exercises has been thoroughly classroom tested for a number of years now.

The book has now a collection of almost six hundred examples, problems, and exercises. Every chapter contains: (a) a number of solved examples each of which is designed to illustrate a specific concept pertaining to a particular section within the chapter, (b) plenty of fully solved problems (which come at the end of every chapter) that are generally comprehensive and, hence, cover several concepts at once, and (c) an abundance of unsolved exercises intended for homework assignments. Through this rich collection of examples, problems, and exercises, I want to empower the student to become an independent learner and an adept practitioner of quantum mechanics. Being able to solve problems is an unfailing evidence of a real understanding of the subject.

The second edition is backed by useful resources designed for instructors adopting the book (please contact the author or Wiley to receive these free resources).

The material in this book is suitable for three semesters—a two-semester undergraduate course and a one-semester graduate course. A pertinent question arises: How to actually use the book in an undergraduate or graduate course(s)? There is no simple answer to this question as this depends on the background of the students and on the nature of the course(s) at hand. First, I want to underscore this important observation: As the book offers an abundance of information, every instructor should certainly select the topics that will be most relevant to her/his students; going systematically over all the sections of a particular chapter (notably Chapter 2), one might run the risk of getting bogged down and, hence, ending up spending too much time on technical topics. Instead, one should be highly selective. For instance, for a one-semester course where the students have not taken modern physics before, I would recommend to cover these topics: Sections 1.1–1.6; 2.2.2, 2.2.4, 2.3, 2.4.1–2.4.8, 2.5.1, 2.5.3, 2.6.1–2.6.2, 2.7; 3.2–3.6; 4.3–4.8; 5.2–5.4, 5.6–5.7; and 6.2–6.4. However, if the students have taken modern physics before, I would skip Chapter 1 altogether and would deal with

these sections: 2.2.2, 2.2.4, 2.3, 2.4.1–2.4.8, 2.5.1, 2.5.3, 2.6.1–2.6.2, 2.7; 3.2–3.6; 4.3–4.8; 5.2–5.4, 5.6–5.7; 6.2–6.4; 9.2.1–9.2.2, 9.3, and 9.4. For a two-semester course, I think the instructor has plenty of time and flexibility to maneuver and select the topics that would be most suitable for her/his students; in this case, I would certainly include some topics from Chapters 7–11 as well (but not all sections of these chapters as this would be unrealistically time demanding). On the other hand, for a one-semester graduate course, I would cover topics such as Sections 1.7–1.8; 2.4.9, 2.6.3–2.6.5; 3.7–3.8; 4.9; and most topics of Chapters 7–11.

Acknowledgments

I have received very useful feedback from many users of the first edition; I am deeply grateful and thankful to everyone of them. I would like to thank in particular Richard Lebed (Arizona State University) who has worked selflessly and tirelessly to provide me with valuable comments, corrections, and suggestions. I want also to thank Jearl Walker (Cleveland State University)—the author of *The Flying Circus of Physics* and of the Halliday–Resnick–Walker classics, *Fundamentals of Physics*—for having read the manuscript and for his wise suggestions; Milton Cha (University of Hawaii System) for having proofread the entire book; Felix Chen (Powerwave Technologies, Santa Ana) for his reading of the first 6 chapters. My special thanks are also due to the following courteous users/readers who have provided me with lists of typos/errors they have detected in the first edition: Thomas Sayetta (East Carolina University), Moritz Braun (University of South Africa, Pretoria), David Berkowitz (California State University at Northridge), John Douglas Hey (University of KwaZulu-Natal, Durban, South Africa), Richard Arthur Dudley (University of Calgary, Canada), Andrea Durlo (founder of the A.I.F. (Italian Association for Physics Teaching), Ferrara, Italy), and Rick Miranda (Netherlands). My deep sense of gratitude goes to M. Bulut (University of Alabama at Birmingham) and to Heiner Mueller-Krumbhaar (Forschungszentrum Juelich, Germany) and his Ph.D. student C. Gugenberger for having written and tested the C++ code listed in Appendix C, which is designed to solve the Schrödinger equation for a one-dimensional harmonic oscillator and for an infinite square-well potential.

Finally, I want to thank my editors, Dr. Andy Slade, Celia Carden, and Alexandra Carrick, for their consistent hard work and friendly support throughout the course of this project.

N. Zettili
Jacksonville State University, USA
January 2009

Preface to the First Edition

Books on quantum mechanics can be grouped into two main categories: textbooks, where the focus is on the formalism, and purely problem-solving books, where the emphasis is on applications. While many fine textbooks on quantum mechanics exist, problem-solving books are far fewer. It is not my intention to merely add a text to either of these two lists. My intention is to combine the two formats into a single text which includes the ingredients of both a textbook and a problem-solving book. Books in this format are practically nonexistent. I have found this idea particularly useful, for it gives the student easy and quick access not only to the essential elements of the theory but also to its practical aspects in a unified setting.

During many years of teaching quantum mechanics, I have noticed that students generally find it easier to learn its underlying ideas than to handle the practical aspects of the formalism. Not knowing how to calculate and extract numbers out of the formalism, one misses the full power and utility of the theory. Mastering the techniques of problem-solving is an essential part of learning physics. To address this issue, the problems solved in this text are designed to teach the student how to calculate. No real mastery of quantum mechanics can be achieved without learning how to derive and calculate quantities.

In this book I want to achieve a double aim: to give a self-contained, yet concise, presentation of most issues of nonrelativistic quantum mechanics, and to offer a rich collection of fully solved examples and problems. This unified format is not without cost. Size! Judicious care has been exercised to achieve conciseness without compromising coherence and completeness.

This book is an outgrowth of undergraduate and graduate lecture notes I have been supplying to my students for about one decade; the problems included have been culled from a large collection of homework and exam exercises I have been assigning to the students. It is intended for senior undergraduate and first-year graduate students. The material in this book could be covered in three semesters: Chapters 1 to 5 (excluding Section 3.7) in a one-semester undergraduate course; Chapter 6, Section 7.3, Chapter 8, Section 9.2 (excluding fine structure and the anomalous Zeeman effect), and Sections 11.1 to 11.3 in the second semester; and the rest of the book in a one-semester graduate course.

The book begins with the experimental basis of quantum mechanics, where we look at those atomic and subatomic phenomena which confirm the failure of classical physics at the microscopic scale and establish the need for a new approach. Then come the mathematical tools of quantum mechanics such as linear spaces, operator algebra, matrix mechanics, and eigenvalue problems; all these are treated by means of Dirac's bra-ket notation. After that we discuss the formal foundations of quantum mechanics and then deal with the exact solutions of the Schrödinger equation when applied to one-dimensional and three-dimensional problems. We then look at the stationary and the time-dependent approximation methods and, finally, present the theory of scattering.

I would like to thank Professors Ismail Zahed (University of New York at Stony Brook) and Gerry O. Sullivan (University College Dublin, Ireland) for their meticulous reading and comments on an early draft of the manuscript. I am grateful to the four anonymous reviewers who provided insightful comments and suggestions. Special thanks go to my editor, Dr Andy Slade, for his constant support, encouragement, and efficient supervision of this project.

I want to acknowledge the hospitality of the Center for Theoretical Physics of MIT, Cambridge, for the two years I spent there as a visitor. I would like to thank in particular Professors Alan Guth, Robert Jaffee, and John Negele for their support.

Note to the student

We are what we repeatedly do. Excellence, then, is not an act, but a habit.
Aristotle

No one expects to learn swimming without getting wet. Nor does anyone expect to learn it by merely reading books or by watching others swim. Swimming cannot be learned without practice. There is absolutely no substitute for throwing yourself into water and training for weeks, or even months, till the exercise becomes a smooth reflex.

Similarly, physics *cannot be learned passively*. Without tackling various challenging problems, the student has no other way of testing the quality of his or her understanding of the subject. Here is where the student gains the sense of satisfaction and involvement produced by a genuine understanding of the underlying principles. *The ability to solve problems is the best proof of mastering the subject*. As in swimming, the more you solve problems, the more you sharpen and fine-tune your problem-solving skills.

To derive full benefit from the examples and problems solved in the text, avoid consulting the solution too early. If you cannot solve the problem after your first attempt, try again! If you look up the solution only after several attempts, it will remain etched in your mind for a long time. But if you manage to solve the problem on your own, you should still compare your solution with the book's solution. You might find a shorter or more elegant approach.

One important observation: as the book is laden with a rich collection of fully solved examples and problems, one should absolutely avoid the temptation of memorizing the various techniques and solutions; instead, one should focus on understanding the concepts and the underpinnings of the formalism involved. It is not my intention in this book to teach the student a number of tricks or techniques for acquiring good grades in quantum mechanics classes without genuine understanding or mastery of the subject; that is, I didn't mean to teach the student how to pass quantum mechanics exams without a deep and lasting understanding. However, the student who focuses on understanding the underlying foundations of the subject and on reinforcing that by solving numerous problems and thoroughly understanding them will doubtlessly achieve a double aim: reaping good grades as well as obtaining a sound and long-lasting education.

N. Zettili

Chapter 1

Origins of Quantum Physics

In this chapter we are going to review the main physical ideas and experimental discoveries that defied classical physics and led to the birth of modern physics, most notably quantum mechanics. The introduction of quantum mechanics was prompted by the failure of classical physics in explaining a number of microphysical phenomena that were observed in the second half of the nineteenth century.

1.1 Historical Note: From Classical to Modern Physics

At the end of the nineteenth century, physics consisted essentially of classical mechanics, the theory of electromagnetism[1], and thermodynamics and statistical mechanics; the body of these broad areas of physics later became known as *classical physics*. Classical mechanics was used to predict the dynamics of *material bodies*, and Maxwell's electromagnetism provided the proper framework to study *radiation*; *matter* and *radiation* were described in terms of *particles* and *waves*, respectively. As for the interactions between matter and radiation, they were well explained by the Lorentz force or by thermodynamics. Classical physics is known for some of its most sacrosanct principles, most notably its *deterministic* nature and the way it treats *particles and waves as mutually exclusive*. Knowing the initial state of a system or object, classical mechanics can provide us with an accurate (i.e., deterministic) description of the system at any later time; for instance, one can determine jointly and with *infinite accuracy* the position and speed of the object and, hence, its trajectory at any time. As a result, the overwhelming success of classical physics in the nineteenth century made people (naively) believe that the ultimate description of nature had been achieved. It seemed that all known physical phenomena could be explained within the framework of the general theories of matter and radiation that were available prior to the middle of the nineteenth century.

Then, between 1859 to 1900, classical physics was found to suffer from a number of untractable limitations: a number of newly discovered phenomena turned out to defy classical physics for over four decades. After a number of failures in attempting to overcome these limitations, it was recognized that one had to go beyond classical physics. As we see below, these efforts have led to the birth of quantum mechanics.

[1]Maxwell's theory of electromagnetism had unified the, then ostensibly different, three branches of physics: electricity, magnetism, and optics.

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

So in the rest of this section, we are going to: (a) briefly discuss some of the well known discoveries that have defied classical physics between 1859–1900 and (b) go over the main ideas and breakthroughs that have led to the birth of quantum mechanics.

1.1.1 1859–1900: Failure of Classical Physics at the Microscopic Scale

In the second half of the nineteenth century, and as soon as experimental techniques were developed to the point of probing deep into matter at the microscopic scale (i.e., at atomic and subatomic structures), a number of new phenomena that challenged classical physics were discovered; no scientist was able to explain them within the framework of classical physics, notwithstanding its previous vast success. Chiefly among these new discoveries that defied classical physics, we may mention four edifying cases: the *blackbody problem, the Balmer and Rydberg formulas, the photoelectric effect, and the Zeeman effect.*

1859: Blackbody radiation Problem
Through a series of experiments, Kirchhoff[2] has shown that the absorption or emission spectrum from a *blackbody*[3] depends on the frequencies of the radiation and on the temperature. Between 1860 and 1900, physicists have struggled to no avail to explain the blackbody problem within the context of classical physics; the most serious attempts, that yet ended in failure, were due to J. Stefan[4] in 1879 and his student Boltzman[5] in 1884, Wien[6] in 1896, and Rayleigh[7] in 1900.

We will see in Subsection 1.1.2 that Planck has managed to solve the blackbody problem in 1900 only after breaking away from classical physics by introducing a novel, paradigmatic idea that ushered in a new era in physics.

1885 and 1888: Balmer and Rydberg Formulas
After years of trial-and-error while studying the light emitted by a hydrogen discharge source (in a discharge or an incandescent lamp), Balmer[8], a Swiss high-school teacher, managed in 1885 to obtain a numerical relationship between the wavelengths of the *visible* spectral lines:

$$\frac{1}{\lambda} = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right), \quad n \text{ is an integer with } n > 2, \quad \text{and} \quad R = 1.0974 \times 10^7\,\mathrm{m}^{-1}, \tag{1.1}$$

where $n = 3, 4, 5, \cdots, \infty$. Later on, these lines became known as the Balmer series and R as the Rydberg constant.

In 1888, Rydberg[9] has modified the Balmer series formula to include all spectral series for

---

[2]G. Kirchhoff, Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin, 662 (1859); 783 (1859); Annalen der Physik und Chemie 109 (2), 275 (1860); Philosophical Magazine, Series 4, Vol. 20, 1 (1860).

[3]A blackbody is an idealized object (e.g., a metallic material) that absorbs all radiation falling on it, regardless of of the frequency or angle of incidence of the radiation. Being a perfect absorber of all incident radiation falling on it, this object will appear "black." We will study in more details this topic in Subsection 1.2.1.

[4]Joseph Stefan, Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, 79, pp. 391—428 (1879).

[5]Ludwig Boltzmann, Ludwig, Annalen der Physik und Chemie 258 (6), pp. 291–294 (1884).

[6]Wilhelm Wien, Annalen der Physik 294, pp. 662–669 (1896).

[7]Lord Rayleigh, Philosophical Magazine 49, pp. 539–540 (1900).

[8]J. J. Balmer, Annalen der Physik und Chemie 25, pp. 80–85 (1885).

[9]J. R. Rydberg, Kongliga Svenska Vetenskaps-Akademiens Handlingar, 23 (11), pp. 1–177. English summary: J. R. Rydberg, *"On the structure of the line-spectra of the chemical elements,"* Philosophical Magazine, 5th series 29, pp. 331–337 (1890).

Table 1.1: Series corresponding to the various spectral emission lines of hydrogen.

n_1	n_2	Series Name	Radiation Type
1	2, 3, 4, $\cdots$, ∞	Lyman series	ultraviolet
2	3, 4, 5, $\cdots$, ∞	Balmer series	visible
3	4, 5, 6, $\cdots$, ∞	Paschen series	infrared
4	5, 6, 7, $\cdots$, ∞	Brackett series	far-infrared
5	6, 7, 8, $\cdots$, ∞	Pfund series	far-infrared
6	7, 8, 9, $\cdots$, ∞	Humphreys series	far-infrared

the hydrogen atom and obtained a more general relation:

$$\frac{1}{\lambda} = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right), \qquad n_1 \text{ and } n_2 \text{ are integers with } n_2 > n_1. \tag{1.2}$$

Later on, this relation became known as the Rydberg formula. Notice that, when $n_1 = 2$, the Rydberg formula (1.2) reduces to the Balmer equation (1.1).

It turned out that the two formulas (1.1) and (1.2) agree well with the experimental data. In spite of the fact that both of them accurately reproduce the experimental results, neither Balmer nor Rydberg knew why their formulas worked. This was due to the fact that Balmer and Rydberg have not obtained them from first principles; they derived them in an ad hoc manner using sheer mathematical insight while trying to numerically reproduce the experimental data. In hindsight, this is expected since classical physics lacks the tools to deal with phenomena taking place at the microscopic scale such as atomic spectroscopy.

As we will see in Subsection 1.1.2, we had to wait until 1911 when Niels Bohr has successfully obtained these formulas by using a mixture of classical physics as well as Planck's new idea (see Eq. (1.85)).

Remark

The Balmer formula (1.1) describes only the Balmer series spectral lines which correspond to **visible** light. However, the Rydberg formula (1.2) is more general and describes all series, including Balmer's, and covers all wavelengths ranging from far-infrared to ultraviolet as illustrated in Table 1.1.

1887: Photoelectric Effect Problem

In a historical experiment, Hertz[10] has discovered the *photoelectric effect* in 1887. In this experiment, Hertz observed that electrons get dislodged and ejected from a metal after shining it with ultra-violet radiation. For a period of almost two decades, none was able to solve this problem within the context of classical physics.

It was only after seeking a solution outside classical physics, by relying on Planck's new idea, that Einstein has managed to find a satisfactory explanation to the photoelectric problem in 1905 (see Subsection 1.1.2).

1896: Zeeman Effect

In his studies of the effect produced by a strong magnetic field on the light emitted by a

[10] H. Hertz, Annalen der Physik **267 (8)**, pp. 983–1000 (1887).

radiant object, Pieter Zeeman[11] has observed in 1896 the *splitting of a spectral line into several components* (of varying frequencies) as light enters a magnetic field. This effect later became known as the *Zeeman effect*.

Since the Zeeman effect takes place at the atomic scale, since the only theoretical tools available to scientists at that time (1896) was classical physics, and since classical physics does not deal with atomic nor with subatomic scales, the Zeeman effect remained an open question for several decades. We will see in Chapter 5 (Subsection 5.6.1) how the Zeeman effect got a satisfactory explanation, but not until it was found outside the scope of classical physics.

Nineteenth Century ended with several open questions
In summary, the nineteenth century ended with several open questions due to the limitations of Classical Physics. These include the phenomena outlined above – the blackbody problem, the photoelectric effect, the Balmer and Rydberg formulas, and the Zeeman effect – as well as others such as the origin of the Mendeleev table that was developed in 1869 by Mendeleev[12]; the X-rays that were discovered by Röntgen[13] in 1895; radioactivity that was accidentally discovery by Becquerel[14] in 1896, etc. None was able to explain why the Mendeleev table had that structure (as we will see in Chapter 8, one can easily obtain the Mendeleev table from first principles using quantum mechanics), what is the origin of the X-rays, and what is the underlying physics involved in radioactivity.

These open questions that have remained unanswered for several decades have starkly demonstrated and confirmed that *classical physics*, which had been quite unassailable, *ceases to be valid at the microscopic scale*. It then became evident that there was no way out except by seeking a *paradigm shift – a major breakaway from classical physics*.

In the rest of this section, we are going to give a brief account of the most consequential ideas that were innovated and developed to successfully solve the outstanding problems mentioned above and how these ideas have led to the birth of quantum mechanics.

1.1.2 1900–1932: Advent of Quantum Mechanics – New Era in Physics

1900: Blackbody Problem Solved Using the Novel Concept of Radiation Quantization
After almost four decades of successive failed attempts by scientists to find solutions to the outstanding problems of that time within the context of classical physics, Max Planck[15] has finally succeeded in pinning down the blackbody[16] problem in 1900, but not until he had sacrificed one its most sacrosanct principles: the *continuous* nature of radiation. As we will see in Subsection 1.2.1, Planck postulated that the energy exchange between radiation and matter takes place in *discrete* or *quantum* amounts and not in continuous amounts as classical

[11] P. Zeeman, Reports of the Ordinary Sessions of the Mathematical and Physical Section (Royal Academy of Sciences in Amsterdam), 5, pp. 181–184 and 242–248 (1896); P. Zeeman, Philosophical Magazine 43 (262), pp. 226–239 (1897); P. Zeeman, Nature, 55 (1424), 347 (1897).

[12] D. Mendeleev, Journal für Praktische Chemie 106, 251 (1869).

[13] W. Röntgen, Aus den Sitzungsberichten der Würzburger Physik.-medic. Gesellschaft Würzburg, pp. 137–147 (1895); pp. 11–17 (1896); Mathematische und Naturwissenschaftliche Mitteilungen aus den Sitzungsberichten der Königlich Preußischen Akademie der Wissenschaften zu Berlin, pp. 392–406 (1897).

[14] H. Becquerel, Comptes Rendus, 122, pp.501–503 (1896).

[15] M. Planck, Verhandlungen der Deutschen Physikalischen Gesellschaft 2, 202 (1900).

[16] Two names are associated with the blackbody problem, Chirchhoff who discovered it and Plank who explained it. While Planck was a junior professor at the university of Berlin, Kirchhoff was chairman of theoretical physics there; after Kirchhoff's death in 1887, Planck was able to succeed him in the chairmanship position.

physics dictates. Planck argued that the energy exchange between an *electromagnetic wave* of frequency ν and matter occurs *only in integer multiples* of $h\nu$,

$$E_n = nh\nu, \qquad n = 0, 1, 2, 3, \cdots, \infty, \tag{1.3}$$

with $E_1 = h\nu$ is the energy of one *quantum* of electromagnetic radiation, where h is a fundamental constant called *Planck's constant*. After obtaining his radiation law, Planck was able to reproduce with great accuracy the spectral density of electromagnetic radiation emitted by a blackbody.

It is worth mentioning that, in addition to representing a major breakaway from classical physics, Planck's idea has ushered in a new era in physics: the birth of *quantum physics*.

Besides giving an accurate description of blackbody radiation, the new *quantum* concept of Planck turned out to be a *paradigmatic idea with far-reaching consequences*. It prompted new thinking and triggered an avalanche of new discoveries that yielded solutions to the most outstanding problems of the time.

Einstein, Bohr, de Broglie, Heisenberg, Schrödinger, Pauli, and Dirac were among those first scientists who took seriously Planck's new ***quantum*** idea, provided it with powerful consolidation, and harnessed it for their respective historical discoveries; their combined contributions have changed and shaped physics for more than a century now. As we will see below, the work of Planck was so ground-breaking and consequential that it inspired and enabled Einstein to explain the photoelectric effect, Bohr to develop a theory of atomic structure that successfully describes atomic spectra (including the derivation of the Balmer and Rydberg formulas), de Broglie to obtain his famous wavelength-momentum equation which shows that particles display wavelike properties, Pauli to propose the exclusion principle that explains the shell structure of atoms and the variety of chemical elements and their combinations, Heisenberg to develop the matrix formulation of quantum mechanics, Schrödinger to propose the wave formulation of quantum mechanics, and Dirac to extend the theory of quantum mechanics to the relativistic domain and to successfully predict the existence of antimatter.

1905: Photoelectric effect problem solved using the concept of the light quantum
In trying to explain the photoelectric effect, Einstein[17] recognized that Planck's idea of the *quantization of the electromagnetic* waves must be valid for *light* as well. As a result, he posited that *light itself is made of discrete bits of energy (or tiny particles)*, called *photons*, each of energy $h\nu$,

$$E_{photon} = h\nu, \tag{1.4}$$

where ν is the frequency of the light. The introduction of the photon concept enabled Einstein to give an elegantly accurate explanation to the photoelectric effect, which had been waiting for a solution ever since its first experimental observation by Hertz in 1887. We will present a quantitative treatment of the photoelectric effect in Subsection 1.2.2.

1905: Einstein introduces the special theory of relativity
In 1905, Einstein[18] has introduced major changes in classical mechanics to account for objects moving at speeds comparable to the speed of light. As a result, he managed to reconcile

[17] A. Einstein, Annalen der Physik **17** (6), pp. 132–148 (1905).
[18] A. Einstein, Annalen der Physik **17** (10), pp. 891–921 (1905).

classical mechanics with Maxwell's theory of electromagnetism[19]. This theory was used in 1928 by Dirac to extend quantum mechanics to the relativistic domain.

1913: Bohr proposed his quantized shell model of the atom and obtained the Balmer-Rydberg formula

Another seminal breakthrough was due to Niels Bohr. Right after the experimental discovery of the atomic nucleus by Rutherford [20] in 1911, and combining Rutherford's atomic model, Planck's quantum concept, and Einstein's photons, Bohr[21] introduced in 1913 his model of the hydrogen atom which consisted of a discrete number of shells (or orbits); the model rests on three postulates: (1) the electron moves around the nucleus in a circular orbit and that only a *discrete set* of classically allowed orbits are available to the electron; (2) the electron's angular momentum in the orbit is *quantized*, $L = n\hbar$ where $\hbar = h/2\pi$ and $n = 1, 2, 3, \cdots$; and (3) atoms can be found only in *discrete energy states* and that the change in an electron's energy results from the electron making *quantum jumps* from one orbit to another and that each jump (to a higher or lower energy orbit) is always accompanied by the emission or absorption of a photon of the right frequency ν and the right energy $h\nu$. Bohr argued that the interaction of atoms with radiation (i.e., the emission or absorption of radiation (photons) by atoms) takes place only in *discrete amounts* of $h\nu$.

As we will see in Subsection 1.6.2.2, the Bohr model provided accurate description and satisfactory explanation to several outstanding problems such as atomic stability and atomic spectroscopy including the reproduction of the derivation of the Balmer-Rydberg formulas from first principles. Clearly, the Bohr model results from a blend of classical physics and quantum physics; it is a semi-classical model because it combines the classical concept of an electron's orbit (Postulate 1) with the quantization concepts introduced by Plank and Einstein (Postulates 2 and 3).

1922: Experimental Evidence for the Quantization of Angular Momentum

The first experimental evidence for the *quantization* of the spatial orientation of angular momentum at the atomic scale was provided by Stern and Gerlach[22] in 1922 using silver (Ag) atoms. As we will see shortly below as well as in Subsection 5.6.1, although the concept of *"spin"* (the intrinsic angular momentum) of the electron was postulated by Goudsmith and Uhlenbeck in 1925, the Stern-Gerlach experiment provided the first direct evidence on the existence of the electron's spin and that it is quantized.

1923: Confirmation that photons are particles (quantum aspect of radiation)

The discovery of Compton[23] in 1923 provided the most conclusive evidence for the corpuscular aspect of radiation. By studying the scattering X-rays with electrons, he confirmed that the X-ray photons behave like particles with momenta $h\nu/c$, where ν is the frequency of the X-rays. "This remarkable agreement between our formulas and the experiments can leave but little doubt that the scattering of X-rays is a quantum phenomenon," Compton concluded in his paper.

---

[19] J. C. Maxwell, Philosophical Transactions of the Royal Society of London 155, pp. 459—512 (1865); J. C. Maxwell, *A Treatise on Electricity and Magnetism*, Volumes 1 and 2, Clarendon Press, Oxford, 1873.

[20] E. Rutherford, Philosophical Magazine, Series 6, vol. 21, 669 (1911).

[21] N. Bohr, Philosophical Magazine, 21 (151), 1 (1913); 26 (153), 476 (1913); 26 (155), 857 (1913).

[22] W. Gerlach and O. Stern, Zeitschrift für Physik, 9, pp. 349–352 (1922); W. Gerlach and O. Stern, Zeitschrift für Physik, 9, pp. 353–355 (1922).

[23] A. Compton, Phys. Rev. 21, pp. 483-502 (1923).

1924: de Broglie's hypothesis about the particle aspect of waves
As if things were not bad enough for classical physics, de Broglie[24], as part of his doctoral thesis, introduced in 1924 another novel idea that classical physics has no theoretical tool to reconcile with: he postulated that not only does radiation exhibit particle-like behavior but, conversely, *material particles* themselves display *wave-like* behavior. He predicted that an electron of momentum p can behave as a wave of wavelength λ (very much like X-rays), where p and λ are connected by $p = h/\lambda$, which is known as the de Broglie relation. This wavelike behavior of electrons was confirmed experimentally in 1927 by Davisson and Germer[25]. They conducted a series of experiments in which they showed that diffraction and interference patterns can be obtained with material particles such as electrons. According to classical physics, diffraction and interference patterns are properties of waves only; these patterns cannot be obtained from particles.

1925: Pauli exclusion principle
In the early part of the twentieth century, it was recognized that atoms with even numbers of electrons are more chemically stable than those with odd numbers of electrons. In 1919, Langmuir[26] suggested that the periodic table could be explained if electrons in an atom occupy a set of shells around the nucleus. Then in 1921, Bohr[27] has revised his model of the atom by suggesting that certain numbers of electrons such as 2, 8, and 18 corresponded to *closed shells* and that these shells are *stable*. In his atomic model, Bohr described the electron's quantum state by means of three quantum numbers which correspond to the electron's three spatial degrees of freedom due to its orbital motion around the nucleus. In 1924, Pauli proposed to add a fourth quantum number to the three numbers that were introduced earlier by Bohr; he argued that the fourth quantum number can take only two values and that it has no classical counterpart. As we will see shortly below, Goudsmith and Uhlenbeck were the first to suggest the idea that this fourth quantum number corresponds to the *spin* angular momentum and that is has nothing to do with the electron's spatial degrees of freedom. Using this fourth quantum number, Pauli was able to explain the closed shell problem that was suggested by Bohr. Then in 1925, Pauli[28] went one step further by introducing the *exclusion principle* which stated that *no two electrons can occupy the same quantum state simulateneously in an atom*; that is, no two electrons in an atom can occupy a state with the same values for the four quantum numbers. This principle, which later became known as the *Pauli Exclusion Principle*, has allowed Pauli to finally explain the structure of the periodic table. In 1940, Pauli has generalized the exclusion principle to all fermions (i.e., half-integer–spin particles) with the *spin-statistics theorem* which implies that all *half-integer–spin particles obey the exclusion principle, while integer-spin particles (bosons) do not.*

1925: Discovery of the electron spin and explanation of the Zeeman effect
Based on their 1925 analysis of atomic spectra (especially the Zeeman effect) and inspired by Pauli's work, two young graduate students, Goudsmith and Uhlenbeck[29], published a series

[24] L. de Broglie, *Annales de Physique* 3, 22 (1925).

[25] C. J. Davisson and L. H. Germer, Phys. Rev. 30, 705 (1927); Nature 119, 558 (1927).

[26] I. Langmuir, *The Arrangement of Electrons in Atoms and Molecules*, Journal of the American Chemical Society 41 (6), pp. 868–934 (1919).

[27] N. Bohr, Nature 107 (2682), pp. 104–107 (1921).

[28] W. Pauli, Zeitschrift für Physik 31 (1), pp. 765–783 (1925); Z. Physik 32, 794 (1925).

[29] S. Goudsmit and G.E. Uhlenbeck, Physica 5, 266 (1925); G.E. Uhlenbeck and S. Goudsmit, Naturwissenschaften 47, 953 (1925); G.E. Uhlenbeck and S. Goudsmit, Nature 117, 264 (1926); S. Goudsmit and G.E. Uhlenbeck, Physica 6,

of articles in which they postulated that electrons have *spin* (i.e., intrinsic angular momentum). Actually, they have simply interpreted the fourth quantum number that was introduced by Pauli as the electron's spin. Although the concept of spin was introduced explicitly at the later date of 1925 by Goudsmith and Uhlenbeck, the first direct evidence on the existence of the electron's spin was actually observed three years earlier by the Stern-Gerlach experiment (see Subsection 5.6.1). However, the Stern-Gerlach experiment was not designed to study the spin (a concept that was not yet introduced in 1922), it was intended to test the Bohr-Sommerfeld hypothesis about the *spatial* quantization of the angular momentum at the atomic scale. Additionally, Pauli came very close to discovering the electron's spin when he proposed the idea that the electron has a fourth quantum number that can take only two values. From a historical perspective, it is evident that Stern and Gerlach in 1922 and Pauli in 1925 have narrowly missed the chance to be the first ones to recognize the existence of the electron's spin.

After introducing the spin concept in 1925, Goudsmith and Uhlenbeck were able to finally provide a satisfactory explanation for the Zeeman effect that has been waiting for an explanation since its discovery in 1896; we will deal with this topic in Subsection 5.6.1.

1925–1926: Two formulations of quantum mechanics

Between 1900–1925, quantum theory was shaped by the eclectic contributions of giants such as Planck, Einstein, Bohr, Sommerfeld, de Broglie, Jordan, Born, Pauli, and others. The collection of these contributions later became known as the *old quantum theory*; this is to contrast it with *modern quantum mechanics* that was introduced in 1925 as we will see shortly below. This period was dominated by hand-waving reasoning and patch-work that was aimed predominantly at addressing the various deficiencies of classical physics. For instance, the "quantization" scheme introduced by Planck in 1900 and the postulates adopted by Bohr in building his atomic model (the most important component of the old quantum theory) were quite arbitrary and did not follow from the first principles of a consistent theory.

It was the very dissatisfaction with the arbitrary nature of the ideas of the old quantum theory as well as the need to blend them within the context of a consistent theory that had prompted Heisenberg and Schrödinger to search for the theoretical foundation underlying these new ideas. By 1925 their efforts paid off: they skillfully welded the various experimental findings as well as Bohr's theory of atomic structure into a refined theory: *quantum mechanics*. Historically, there were two independent formulations of quantum mechanics that were proposed almost simultaneously by Heinsenberg and Schrödinger.

1925: Heisenberg's matrix formulation of quantum mechanics

The first formulation, called **matrix mechanics**, was developed by Heisenberg[30] to describe atomic structure starting from the observed *discrete* nature of the spectral lines. Inspired by Planck's quantization of waves and by Bohr's model of the hydrogen atom, Heisenberg founded his theory on the notion that the only allowed values of energy exchange between microphysical systems are those that are discrete: the *quanta* of energy. The central idea of Heisenberg's matrix mechanics is that all physical quantities must be represented by infinite hermitian (i.e., self-adjoint) matrices. In particular, representing the energy with a hermitian matrix (the Hamiltonian), Heisenberg obtained an eigenvalue problem that describes the dynamics of microscopic systems; the diagonalization of the Hamiltonian matrix yields the

273 (1926)

[30]W. Heisenberg, *Zeitschrift für Physik*, **Vol. 33, No. 1**, pp. 879–893 (1925); Die Naturwissenschaften, **14**, pp. 899–894 (1926).

energy spectrum and the state vectors of the system.

Shortly after Heisenberg introduced his theory, Born and Jordan[31] have managed to work out a rigorous formulation of matrix mechanics in 1925; then few months later, Born, Heisenberg, and Jordan[32] have provided a more elaborate treatment of matrix mechanics in 1926.

Heisenberg's matrix mechanics proved very successful in accounting for the empirical results of nearly all spectroscopic data that were available at the time.

1926: Schrödinger's wave formulation of quantum mechanics

The second formulation, called **wave mechanics**, was due to Schrödinger[33] who based it on the de Broglie's idea that matter can have wavelike properties. Dissatisfied with the astronomical analogy ideas that were used by Bohr and Sommerfeld to describe spectral lines in terms of electrons in well-defined elliptical orbits around the nucleus, very much like the orbital motion of planets around the sun, Schrödinger took an entirely different approach. In his theory, which later became known as *wave mechanics*, Schrödinger visualized electrons as wavelike objects with their charges distributed ever space and that the electrons' positions in space can only be found up to a certain relative probability.

The Schrödinger wave formulation, which is more intuitive than matrix mechanics, describes the dynamics of microscopic matter (such as electrons) by means of a *wave equation*, called the *Schrödinger equation*. Instead of the matrix eigenvalue problem of Heisenberg, Schrödinger obtained a *differential equation*. The solutions of this equation yield the energy spectrum and the wave function of the system under consideration.

1926: The matrix and wave formulations of quantum mechanics proved to be equivalent

The two ostensibly different formulations of quantum mechanics by Heisenberg and Schrödinger were shown to be equivalent, mainly by Schrödinger in 1926, Dirac in 1927, and John Von Neuman in 1932 (as will be seen below). Schrödinger[34] was credited to be the first who raised the question in 1926 about the connection between matrix and wave mechanics. Almost simultaneously, Dirac[35] has shown that the Heisenberg and Schrödinger theories can be obtained from a more general formulation of quantum mechanics, called *transformation theory*, that he developed as part of his Ph.D. thesis. In this approach, as we will see in Chapter 2, Dirac used the abstract braket notation where quantum states are described by means of kets (state vectors), the dual transform of kets by bras, and the scalar product between a ket and a bra by a braket; kets and bras are elements of the Hilbert space. The concept of "transformation" refers to the changes that the quantum states undergo in time.

As we will see in Chapter 2, the representation of the Dirac formalism in a *discrete* basis gives back the Heisenberg matrix formulation of quantum mechanics, whereas its representation in a *continuous* basis – the position or momentum representation – gives back Schrödinger's wave mechanics. Dirac refined his general formulation of quantum mechanics and published it in his 1930 treatise[36] that eventually became one of the most consequential manuscripts of quantum mechanics.

---

[31] M. Born and P. Jordan, Zeitschrift für Physik, **34**, pp. 858–888 (1925).

[32] M. Born, W. Heisenberg, and P. Jordan , Zeitschrift für Physik, **35**, pp. 557–615 (1926).

[33] E. Schrödinger, *Ann. Phys.*, **79**, 361 (1926); **79**, 489 (1926); **79**, 734 (1926); **80**, 437 (1926); **81**, 109 (1926); *Die Naturwissenschaften*, **14**, 664 (1926); *Phys. Rev.*, **Vol. 28, No. 6**, 1049 (1926); *Collected Papers on Wave Mechanics*, Blackie & Son, London, 1928.

[34] E. Schrödinger, "On the Relation of the Heisenberg-Born-Jordan Quantum Mechanics and Mine", Annalen der Physik, **79**, pp. 734–756 (1926).

[35] P. A. M. Dirac, Proceedings of the Royal Society of London **A. 113** (765), pp. 621–641 (1927).

[36] P. A. M. Dirac, *The Principles of Quantum Mechanics*, Cambridge University Press (1930).

It is worth mentioning that, while Heisenberg's matrix mechanics seemed quite abstract and not easy to apply to many problems, the shrödinger wave formulation turned out to be practical and easy to understand. Whereas matrix mechanics is found to be effective in only a limited number of problems such as the one-dimensional harmonic oscillator (see Chapter 4), wave mechanics turned out to be very successful in offering solutions to a wide range of problems, ranging from one dimensional to three dimensional problems as we will see in Chapters 4 and 6. One of the great successes of the Schrödinger equation is in providing an accurate description of the hydrogen atom (see Chapter 6).

While the Schrodinger formulation deals with a second-order differential equation (the Schrödinger equation), the Heisenberg approach deals with an abstract eigenvalue problem in a matrix form. In this context, it is worth mentioning that Hilbert had suggested to Heisenberg to find a differential equation that would correspond to his matrix eigenvalue equations and, for some reason, Heisenberg has not acted on this recommendation; had he heeded Hilbert's suggestion, Heisenberg may have discovered the Schrödinger equation before Schrödinger. Due to the limited applicability range of Heisenberg's matrix mechanics and its abstract nature, Schrödinger's wave mechanics eventually became the method of choice of doing quantum mechanics due to its practicality and applicability to wide areas of atomic and molecular physics.

In addition to providing an accurate reproduction of the existing experimental data pertaining to atomic and molecular physics, quantum mechanics turned out to possess an astonishingly reliable predictive power which enabled it to explore and unravel many uncharted areas of the microphysical world. This new theory had put an end to twenty-five years (1900–1925) of ad hoc, patchwork of the old quantum theory.

1926: Probabilistic interpretation of quantum mechanics

Another important contribution to quantum mechanics was due to Max Born[37] in 1926 when he proposed the *probabilistic* interpretation of wave mechanics. He interpreted the square modulus of the wave function (that is solution to the Schrödinger equation) as a *probability density*; he argued that the probability density of finding a particle at a given point is proportional to the square of the magnitude of the particle's wavefunction at that point. This interpretation has successfully established a link between the theory of quantum mechanics and the vast body of experimental results.

The overwhelming success of the probabilistic interpretation of quantum mechanics was viewed as one of the most important turning points in the history of quantum mechanics: the appearance and acceptance of *indeterminism* in physics. Up to the introduction of the Born probabilistic interpretation, physics has been considered to be deterministic; determinism has been one of the most sacrosanct pillars of classical physics. The classical concept of viewing the electron as a pointlike particle moving on a sharply defined orbit around the nucleus is replaced in wave mechanics by a cloud that describes the most probable locations of the electron.

1927: Heisenberg uncertainty principle

In his formulation of matrix mechanics, Heisenberg recognized an important fact: dynamical variables such as position and momentum *do not commute* with each other. This is due to the mathematics of matrices: in general, the product of two matrices A and B do not commute: $AB \neq BA$. For instance, since the matrices representing position and momentum do not

[37] M. Born, Zeitschrift für Physik 38, pp. 803–827 (1926).

commute, Heisenberg has shown that the commutator between the position and momentum matrices is given by $XP - PX = i\hbar$ where $\hbar = h/2\pi$, h being the Planck constant. This commutation relation, which is considered to be the most important commutator of quantum mechanics, represents a major departure from classical physics where dynamical variables commute: $XP = PX$. This observation has led Heisenberg[38] naturally to the *uncertainty principle* which, in its simplest form, states that *we cannot simultaneously measure the position and momentum of a particle with infinite accuracy*; the more precisely one tries to measure the position, the more uncertain the momentum would become, and vice versa. This principle applies also to energy and time: *energy of a particle and time cannot be simultaneously measured with infinite accuracy*. Heisenberg derived the following historically significant mathematical relations governing the uncertainties between position-momentum and energy-time:

$$\Delta x \Delta p \geq \hbar/2, \qquad \Delta E \Delta t \geq \hbar/2, \tag{1.5}$$

where Δx, Δp, ΔE, and Δt are the uncertainties in position, momentum, energy, and time, respectively, and $\hbar = h/2\pi$.

To formulate the uncertainty principle, Heisenberg had conducted a Gedanken (thought) experiment in which he envisioned measuring the position of an electron with a gamma ray microscope. The high-energy photon used to probe the electron would impart a kick on the electron causing its momentum to change in an uncertain way. Hence, a higher resolution microscope, which will gives us a more precise reading of the electron's position, would obviously require a higher energy beam of gamma rays; this would end up imparting an even greater momentum on the electron and, hence, would cause an even greater uncertainty about is momentum. Consequently, the more precisely one tries to measure the position, the more uncertain the momentum would become, and vice versa. Heisenberg argued that this uncertainty is a fundamental property of quantum mechanics and that it is not due to a limitation of any particular experimental apparatus. It is worthy noting that the uncertainty principle constitutes another major breakaway from classical physics which is founded on the premise that, in principle, one can obtain exact values for all physical quantities simultaneously. So the certain or infinitely accurate outcomes of measurements achieved by classical physics at the macroscopic level is something that cannot be achieved at the microscopic scale.

1928: Dirac's theory of relativistic quantum mechanics and its prediction of antimatter
Modern physics rests on two main pillars: relativity and quantum mechanics. The idea of blending these two theories into a single theory – *relativistic quantum mechanics* – became a necessity, at least from an esthetic perspective. This is exactly what Dirac has achieved in 1928. Combining the theory of special relativity with quantum mechanics, Dirac[39] derived an equation which describes the motion of electrons moving at *relativistic speeds*. This equation, known as Dirac's equation, predicted the existence of an antiparticle – the positron – which has similar properties, but opposite charge, with the electron.

1932: Experimental evidence for the existence of antimatter
Several years after its theoretical prediction by Dirac, the positron was discovered experimentally by Anderson[40] in 1932 while studying comic rays. During the course of his experiments

[38] W. Heisenberg, Zeitschrift fúr Physik 43, pp. 172-198 (1927); English translation in *Wheeler and Zurek*, pp. 62-84 (1983).
[39] P.A.M. Dirac, Proc. Roy. Soc. **A117**, 610 (1928).
[40] Carl D. Anderson, *Phys. Rev.*, **43**, 491 (1933).

with cosmic rays, Anderson encountered unexpected particle tracks in his cloud chamber photographs. After careful analysis of these tracks, Anderson concluded that they were produced by a particle that has the same mass as the electron, but with opposite electrical charge. Shortly after this discovery, Anderson provided a more conclusive evidence on the existence of positrons by creating electron-positron pairs by bombarding various materials with gamma rays. This discovery has given solid a confirmation for the validity of Dirac's theory. Later on, Dirac has generalized this principle to all particles: he correctly predicted that every particle possesses an antiparticle (antiprotons, antineutrons, antimuons, and so on). This bold prediction was quite revolutionary and visionary back in 1932, notwithstanding the fact that the existence of antimater is taken for granted today.

1932: Mathematical foundation of quantum mechanics
After Heisenberg introduced matrix mechanics, John von Neumann, an outstanding mathematician who was dissatisfied with the use of matrices, has embarked on developing his own version of quantum mechanics by working out its underlying mathematics. In 1932, he published his work in a seminal book[41] – *The Mathematical Foundations of Quantum Mechanics* – that has been credited for putting quantum mechanics on solid mathematical ground. In his approach, von Neumann introduced an *operator theory*, which later became known as *Neumann's algebras*. In his theory, Von Neumann was able to explain, with a high degree of mathematical rigor, various aspects of quantum mechanics. By working out the formalism of quantum mechanics in terms of Hermitian operators on Hilbert spaces, Von Neumann managed to construct a theory of quantum mechanics that is more general where matrix and wave mechanics are special cases; he is credited for having shown that matrix and wave mechanics are mathematically equivalent.

Quantum mechanics, the main driver of modern physics
The theory of relativity and quantum mechanics are the two main pillars of modern physics. While relativity deals with fastly moving objects, quantum mechanics handles matter at the microscopic scale. Being the *only* valid framework for describing the microphysical world, quantum mechanics has become indispensable to all ares of modern physics such as solid state, molecular, atomic, nuclear, and particle physics, and so on. Not only that, it is also considered to be the foundation to various areas of science (chemistry and biology) and engineering; most notably, the fields of nanoscience and nanotechnology with their systemic applications to science, technology, pharmacology, and medicine.

1.2 Particle Aspect of Radiation

According to classical physics, a particle is characterized by an energy E and a momentum $\vec{p}$, whereas a wave is characterized by an amplitude and a wave vector $\vec{k}$ ($|\vec{k}| = 2\pi/\lambda$) that specifies the direction of propagation of the wave. Particles and waves exhibit entirely different behaviors; for instance, the *particle and wave properties are mutually exclusive*. We should note that waves can exchange *any* (continuous) amount of energy with particles.

In this section we are going to see how these rigid concepts of classical physics led to its failure in explaining a number of microscopic phenomena such as blackbody radiation, the

[41] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, (Springer-Verlag, Berlin, 1932). English translation: *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955).

photoelectric effect, and the Compton effect. As it turned out, these phenomena could only be explained by abandoning the rigid concepts of classical physics and introducing a new concept: the *particle* aspect of radiation.

1.2.1 Blackbody Radiation

At issue here is how radiation interacts with matter. When heated, a solid object glows and emits thermal radiation. As the temperature increases, the object becomes red, then yellow, then white. The thermal radiation emitted by glowing solid objects consists of a *continuous* distribution of frequencies ranging from infrared to ultraviolet. The continuous pattern of the distribution spectrum is in sharp contrast to the radiation emitted by heated gases; the radiation emitted by gases has a discrete distribution spectrum: a few sharp (narrow), colored lines with no light (i.e., darkness) in between.

Understanding the continuous character of the radiation emitted by a glowing solid object constituted one of the major unsolved problems during the second half of the nineteenth century. All attempts to explain this phenomenon by means of the available theories of classical physics (statistical thermodynamics and classical electromagnetic theory) ended up in miserable failure. This problem consisted in essence of specifying the proper theory of thermodynamics that describes how energy gets exchanged between radiation and matter.

When radiation falls on an object, some of it might be absorbed and some reflected. An idealized "blackbody" is a material object that absorbs all of the radiation falling on it, and hence appears as black under reflection when illuminated from outside. When an object is heated, it radiates electromagnetic energy as a result of the thermal agitation of the electrons in its surface. The intensity of this radiation depends on its frequency and on the temperature; the light it emits ranges over the entire spectrum. An object in thermal equilibrium with its surroundings radiates as much energy as it absorbs. It thus follows that a blackbody is a perfect absorber as well as a perfect emitter of radiation.

A practical blackbody can be constructed by taking a hollow cavity whose internal walls perfectly reflect electromagnetic radiation (e.g., metallic walls) and which has a very small hole on its surface. Radiation that enters through the hole will be trapped inside the cavity and gets completely absorbed after successive reflections on the inner surfaces of the cavity. The hole thus absorbs radiation like a black body. On the other hand, when this cavity is heated[42] to a temperature T, the radiation that leaves the hole is blackbody radiation, for the hole behaves as a perfect emitter; as the temperature increases, the hole will eventually begin to glow. To understand the radiation inside the cavity, one needs simply to analyze the spectral distribution of the radiation coming out of the hole. In what follows, the term *blackbody radiation* will then refer to the radiation leaving the hole of a heated hollow cavity; the radiation emitted by a blackbody when hot is called blackbody radiation.

By the mid-1800s, a wealth of experimental data about blackbody radiation was obtained for various objects. All these results show that, at equilibrium, the radiation emitted has a well-defined, continuous energy distribution: to each frequency there corresponds an energy density which depends neither on the chemical composition of the object nor on its shape, but only on the temperature of the cavity's walls (Figure 1.1). The energy density shows a pronounced maximum at a given frequency, which increases with temperature; that is, *the*

[42]When the walls are heated uniformly to a temperature T, they emit radiation (due to thermal agitation or vibrations of the electrons in the metallic walls).

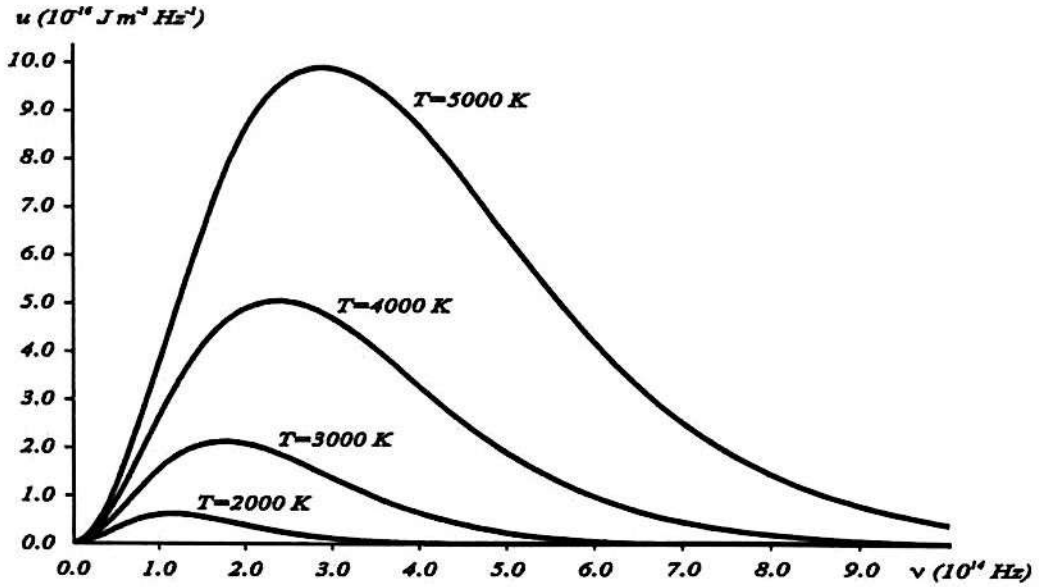


Figure 1.1: Spectral energy density $u(\nu, T)$ of blackbody radiation at different temperatures as a function of the frequency ν.

peak of the radiation spectrum occurs at a frequency that is proportional to the temperature (1.21): $\nu_{max} \propto T$. This is the underlying reason behind the change in color of a heated object as its temperature increases, notably from red to yellow to white. It turned out that the explanation of the blackbody spectrum was not so easy.

A number of attempts aimed at explaining the origin of the continuous character of this radiation were carried out. The most serious among such attempts, and which made use of classical physics, were due to Wilhelm Wien in 1889 and Rayleigh in 1900. In 1879 J. Stefan found *experimentally* that the total intensity (or the total power per unit surface area) radiated by a glowing object of temperature T is given by

$$\mathcal{P} = a\sigma T^4, \tag{1.6}$$

which is known as the Stefan–Boltzmann law, where $\sigma = 5.67 \times 10^{-8}\ \mathrm{W\,m^{-2}\,K^{-4}}$ is the Stefan–Boltzmann constant, and a is a coefficient which is less than or equal to 1; in the case of a blackbody $a = 1$. Then in 1884 Boltzmann provided a *theoretical* derivation for Stefan's experimental law by combining thermodynamics and Maxwell's theory of electromagnetism.

Wien's energy density distribution

Using thermodynamic arguments, Wien took the Stefan–Boltzmann law (1.6) and in 1894 he extended it to obtain the energy density per unit frequency of the emitted blackbody radiation:

$$u(\nu, T) = A\nu^3 e^{-\beta\nu/T}, \tag{1.7}$$

where A and β are empirically defined parameters (they can be adjusted to fit the experimental data). **Note:** $u(\nu, T)$ has the dimensions of an energy per unit volume per unit frequency; its SI units are $\mathrm{J\,m^{-3}\,Hz^{-1}}$. Although Wien's formula fits the high-frequency data remarkably well, it fails badly at low frequencies (Figure 1.2).

Rayleigh's energy density distribution

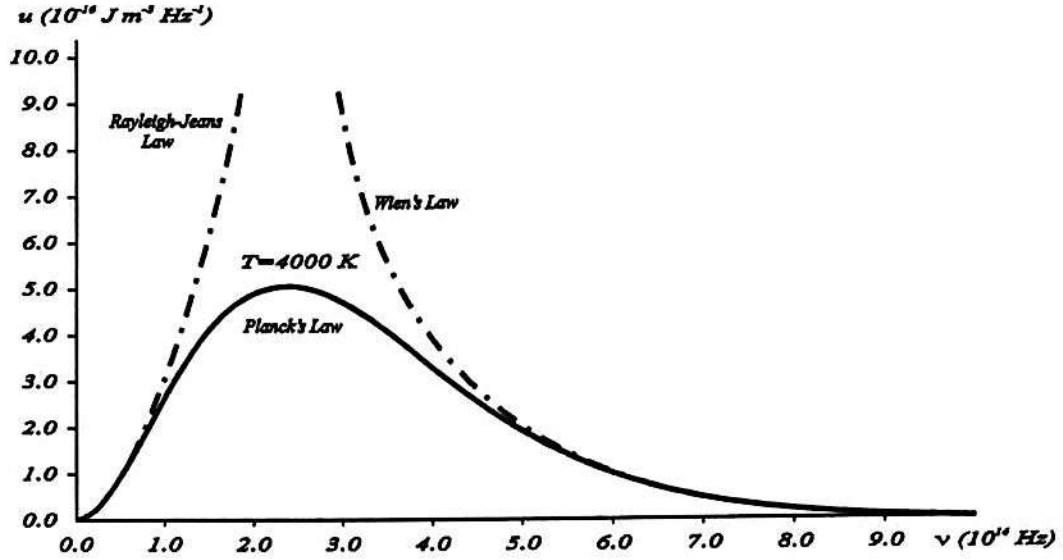


Figure 1.2: Comparison of various spectral densities: while the Planck and experimental distributions match perfectly (solid curve), the Rayleigh–Jeans and the Wien distributions (dotted curves) agree only partially with the experimental distribution.

In his 1900 attempt, Rayleigh focused on understanding the nature of the electromagnetic radiation inside the cavity. He considered the radiation to consist of standing waves having a temperature T with nodes at the metallic surfaces. These standing waves, he argued, are equivalent to harmonic oscillators, for they result from the harmonic oscillations of a large number of electrical charges, electrons, that are present in the walls of the cavity. When the cavity is in thermal equilibrium, the electromagnetic energy density inside the cavity is equal to the energy density of the charged particles in the walls of the cavity; the average total energy of the radiation leaving the cavity can be obtained by multiplying the average energy of the oscillators by the number of modes (standing waves) of the radiation in the frequency interval ν to $\nu + d\nu$:

$$N(\nu) = \frac{8\pi\nu^2}{c^3}, \tag{1.8}$$

where $c = 3 \times 10^8\ \mathrm{m\,s^{-1}}$ is the speed of light; the quantity $(8\pi\nu^2/c^3)d\nu$ gives the number of modes of oscillation per unit volume in the frequency range ν to $\nu + d\nu$. So the electromagnetic energy density in the frequency range ν to $\nu + d\nu$ is given by

$$u(\nu, T) = N(\nu)\langle E\rangle = \frac{8\pi\nu^2}{c^3}\langle E\rangle, \tag{1.9}$$

where $\langle E\rangle$ is the average energy of the oscillators present on the walls of the cavity (or of the electromagnetic radiation in that frequency interval); the temperature dependence of $u(\nu, T)$ is buried in $\langle E\rangle$.

How does one calculate $\langle E\rangle$? According to the equipartition theorem of classical thermodynamics, all oscillators in the cavity have the same mean energy, irrespective of their

frequencies[43]:

$$\langle E\rangle = \frac{\int_0^\infty Ee^{-E/kT}dE}{\int_0^\infty e^{-E/kT}dE} = kT, \tag{1.10}$$

where $k = 1.3807\times10^{-23}\,\mathrm{J\,K^{-1}}$ is the Boltzmann constant. An insertion of (1.10) into (1.9) leads to the Rayleigh–Jeans formula:

$$u(\nu,T) = \frac{8\pi\nu^2}{c^3}kT. \tag{1.11}$$

Except for low frequencies, this law is in complete disagreement with experimental data: $u(\nu,T)$ as given by (1.11) *diverges* for high values of ν, whereas experimentally it must be finite (Figure 1.2). Moreover, if we integrate (1.11) over all frequencies, the integral *diverges*. This implies that the cavity contains an *infinite* amount of energy. This result is absurd. Historically, this was called the *ultraviolet catastrophe*, for (1.11) diverges for *high* frequencies (i.e., in the *ultraviolet* range)—a real catastrophical failure of classical physics indeed! The origin of this failure can be traced to the derivation of the average energy (1.10). It was founded on an erroneous premise: the energy exchange between radiation and matter is *continuous*; any amount of energy can be exchanged.

Planck's energy density distribution

By devising an ingenious scheme—interpolation between Wien's rule and the Rayleigh–Jeans rule—Planck succeeded in 1900 in avoiding the ultraviolet catastrophe and proposed an accurate description of blackbody radiation. In sharp contrast to Rayleigh's assumption that a standing wave can exchange *any* amount (continuum) of energy with matter, Planck considered that the energy exchange between radiation and matter must be *discrete*. He then *postulated* that the energy of the radiation (of frequency ν) emitted by the oscillating charges (from the walls of the cavity) must come *only* in *integer multiples* of $h\nu$:

$$E = nh\nu, \qquad n = 0,\ 1,\ 2,\ 3,\ \cdots, \tag{1.12}$$

where h is a universal constant and $h\nu$ is the energy of a *"quantum"* of radiation (ν represents the frequency of the oscillating charge in the cavity's walls as well as the frequency of the radiation emitted from the walls, because the frequency of the radiation emitted by an oscillating charged particle is equal to the frequency of oscillation of the particle itself). That is, the energy of an oscillator of natural frequency ν (which corresponds to the energy of a charge oscillating with a frequency ν) must be an *integral multiple* of $h\nu$; note that $h\nu$ is not the same for all oscillators, because it depends on the frequency of each oscillator. Classical mechanics, however, puts no restrictions whatsoever on the frequency, and hence on the energy, an oscillator can have. The energy of oscillators, such as pendulums, mass–spring systems, and electric oscillators, varies continuously in terms of the frequency. Equation (1.12) is known as *Planck's quantization rule* for energy or *Planck's postulate*.

So, assuming that the energy of an oscillator is quantized, Planck showed that the *correct* thermodynamic relation for the average energy can be obtained by merely replacing the integration of (1.10)—that corresponds to an energy continuum—by a *discrete* summation

[43] Using a variable change $\beta = 1/(kT)$, we have $\langle E\rangle = -\frac{\partial}{\partial\beta}\ln\left(\int_0^\infty e^{-\beta E}dE\right) = -\frac{\partial}{\partial\beta}\ln(1/\beta) = 1/\beta \equiv kT$.

corresponding to the discreteness of the oscillators' energies[44]:

$$\langle E\rangle = \frac{\sum_{n=0}^{\infty} nh\nu e^{-nh\nu/kT}}{\sum_{n=0}^{\infty} e^{-nh\nu/kT}} = \frac{h\nu}{e^{h\nu/kT}-1}, \tag{1.13}$$

and hence, by inserting (1.13) into (1.9), the energy density per unit frequency of the radiation emitted from the hole of a cavity is given by

$$\boxed{u(\nu,T) = \frac{8\pi\nu^2}{c^3}\frac{h\nu}{e^{h\nu/kT}-1}.} \tag{1.14}$$

This is known as *Planck's distribution*. It gives an exact fit to the various experimental radiation distributions, as displayed in Figure 1.2. The numerical value of h obtained by fitting (1.14) with the experimental data is $h = 6.626 \times 10^{-34}$ J s. We should note that, as shown in Example 1.1 (see Eq. (1.17)), we can rewrite Planck's energy density (1.14) to obtain the energy density per unit wavelength

$$\boxed{\tilde{u}(\lambda,T) = \frac{8\pi hc}{\lambda^5}\frac{1}{e^{hc/\lambda kT}-1}.} \tag{1.15}$$

Let us now look at the behavior of Planck's distribution (1.14) in the limits of both low and high frequencies, and then try to establish its connection to the relations of Rayleigh–Jeans, Stefan–Boltzmann, and Wien. First, in the case of very low frequencies $h\nu \ll kT$, we can show that (1.14) reduces to the Rayleigh–Jeans law (1.11), since $\exp(h\nu/kT) \simeq 1 + h\nu/kT$. Moreover, if we integrate Planck's distribution (1.14) over the whole spectrum (where we use a change of variable $x = h\nu/kT$ and make use of a special integral[45]), we obtain the total energy density which is expressed in terms of Stefan–Boltzmann's total power per unit surface area (1.6) as follows:

$$\int_0^{\infty} u(\nu,T)d\nu = \frac{8\pi h}{c^3}\int_0^{\infty}\frac{\nu^3}{e^{h\nu/kT}-1}d\nu = \frac{8\pi k^4T^4}{h^3c^3}\int_0^{\infty}\frac{x^3}{e^x-1}dx = \frac{8\pi^5k^4}{15h^3c^3}T^4 = \frac{4}{c}\sigma T^4, \tag{1.16}$$

where $\sigma = 2\pi^5k^4/15h^3c^2 = 5.67 \times 10^{-8}\ \mathrm{W\,m^{-2}\,K^{-4}}$ is the Stefan–Boltzmann constant. In this way, Planck's relation (1.14) leads to a *finite* total energy density of the radiation emitted from a blackbody, and hence avoids the ultraviolet catastrophe. Second, in the limit of *high* frequencies, we can easily ascertain that Planck's distribution (1.14) yields Wien's rule (1.7).

In summary, the spectrum of the blackbody radiation reveals the quantization of radiation, notably the particle behavior of electromagnetic waves.

The introduction of the constant h had indeed heralded the end of classical physics and the dawn of a new era: physics of the microphysical world. Stimulated by the success of Planck's quantization of radiation, other physicists, notably Einstein, Compton, de Broglie, and Bohr, skillfully adapted it to explain a host of other outstanding problems that had been unanswered for decades.

---

[44] To derive (1.13) one needs: $1/(1-x) = \sum_{n=0}^{\infty} x^n$ and $x/(1-x)^2 = \sum_{n=0}^{\infty} nx^n$ with $x = e^{-h\nu/kT}$.

[45] In integrating (1.16), we need to make use of this integral: $\int_0^{+\infty}\frac{x^3}{e^x-1}dx = \frac{\pi^4}{15}$.

Example 1.1 (Wien's displacement law)

(a) Show how to Derive Eq. (1.15) which gives the Planck energy density per unit wavelength.

(b) Show that the *maximum* of the Planck energy density per unit wavelength (1.15) occurs for a wavelength of the form $\lambda_{max} = b/T$, where T is the temperature and b is a constant that needs to be estimated.

(c) Use the relation derived in (a) to estimate the surface temperature of a star if the radiation it emits has a maximum intensity at a wavelength of 446 nm. What is the intensity radiated by the star?

(d) Estimate the wavelength and the intensity of the radiation emitted by a glowing tungsten filament whose surface temperature is 3300 K.

Solution

(a) Since $\nu = c/\lambda$, we have $d\nu = |d\nu/(d\lambda)|\, d\lambda = (c/\lambda^2)d\lambda$; we can thus write Planck's energy density (1.14) in terms of the wavelength as follows:

$$\bar{u}(\lambda, T) = u(\nu, T)\left|\frac{d\nu}{d\lambda}\right| = \frac{8\pi hc}{\lambda^5}\frac{1}{e^{hc/\lambda kT} - 1}. \tag{1.17}$$

(b) The *maximum* of $\bar{u}(\lambda, T)$ corresponds to $\partial\bar{u}(\lambda, T)/\partial\lambda = 0$, which yields

$$\frac{8\pi hc}{\lambda^6}\left[-5\left(1 - e^{-hc/\lambda kT}\right) + \frac{hc}{\lambda kT}\right]\frac{e^{hc/\lambda kT}}{\left(e^{hc/\lambda kT} - 1\right)^2} = 0, \tag{1.18}$$

and hence

$$\frac{\alpha}{\lambda} = 5\left(1 - e^{-\alpha/\lambda}\right), \tag{1.19}$$

where $\alpha = hc/(kT)$. We can solve this transcendental equation either graphically or numerically by writing $\alpha/\lambda = 5 - \varepsilon$. Inserting this value into (1.19), we obtain $5 - \varepsilon = 5 - 5e^{-5+\varepsilon}$, which leads to a suggestive approximate solution $\varepsilon \approx 5e^{-5} = 0.0337$ and hence $\alpha/\lambda = 5 - 0.0337 = 4.9663$. Since $\alpha = hc/(kT)$ and using the values $h = 6.626 \times 10^{-34}$ J s and $k = 1.3807 \times 10^{-23}$ J K^{-1}, we can write the wavelength that corresponds to the maximum of the Planck energy density (1.14) as follows:

$$\boxed{\lambda_{max} = \frac{hc}{4.9663k}\frac{1}{T} = \frac{2898.9 \times 10^{-6}\ \text{m K}}{T}.} \tag{1.20}$$

This relation, which shows that λ_{max} decreases with increasing temperature of the body, is called *Wien's displacement law*. It can be used to determine the wavelength corresponding to the maximum intensity if the temperature of the body is known or, conversely, to determine the temperature of the radiating body if the wavelength of greatest intensity is known. This law can be used, in particular, to estimate the temperature of stars (or of glowing objects) from their radiation, as shown in part (b). From (1.20) we obtain

$$\boxed{\nu_{max} = \frac{c}{\lambda_{max}} = \frac{4.9663}{h}kT.} \tag{1.21}$$

This relation shows that the peak of the radiation spectrum occurs at a frequency that is proportional to the temperature.

(c) If the radiation emitted by the star has a maximum intensity at a wavelength of $\lambda_{max} = 446$ nm, its surface temperature is given by

$$T = \frac{2898.9 \times 10^{-6} \text{ m K}}{446 \times 10^{-9} \text{ m}} \simeq 6500 \text{ K}. \tag{1.22}$$

Using Stefan–Boltzmann's law (1.6), and assuming the star to radiate like a blackbody, we can estimate the total power per unit surface area emitted at the surface of the star:

$$\mathcal{P} = \sigma T^4 = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4} \times (6500 \text{ K})^4 \simeq 101.2 \times 10^6 \text{ W m}^{-2}. \tag{1.23}$$

This is an enormous intensity which will decrease as it spreads over space.

(d) The wavelength of greatest intensity of the radiation emitted by a glowing tungsten filament of temperature 3300 K is

$$\lambda_{max} = \frac{2898.9 \times 10^{-6} \text{ m K}}{3300 \text{ K}} \simeq 878.45 \text{ nm}. \tag{1.24}$$

The intensity (or total power per unit surface area) radiated by the filament is given by

$$\mathcal{P} = \sigma T^4 = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4} \times (3300 \text{ K})^4 \simeq 6.7 \times 10^6 \text{ W m}^{-2}. \tag{1.25}$$

1.2.2 Photoelectric Effect

The photoelectric effect provides a direct confirmation for the energy quantization of light. In 1887 Hertz discovered the photoelectric effect: electrons[46] were observed to be ejected from metals when irradiated with light (Figure 1.3a). Moreover, the following experimental laws were discovered prior to 1905:

- If the frequency of the incident radiation is smaller than the metal's threshold frequency—a frequency that depends on the properties of the metal—no electron can be emitted regardless of the radiation's intensity (Philip Lenard, 1902).
- No matter how low the intensity of the incident radiation, electrons will be ejected *instantly* the moment the frequency of the radiation exceeds the threshold frequency ν_0.
- At any frequency above ν_0, the number of electrons ejected increases with the intensity of the light but does not depend on the light's frequency.
- The kinetic energy of the ejected electrons depends on the frequency but not on the intensity of the beam; the kinetic energy of the ejected electron increases *linearly* with the incident frequency.

These experimental findings cannot be explained within the context of a purely classical picture of radiation, notably the dependence of the effect on the threshold frequency. According to classical physics, any (continuous) amount of energy can be exchanged with matter.

[46] In 1899 J. J. Thomson confirmed that the particles giving rise to the photoelectric effect (i.e., the particles ejected from the metals) are electrons.

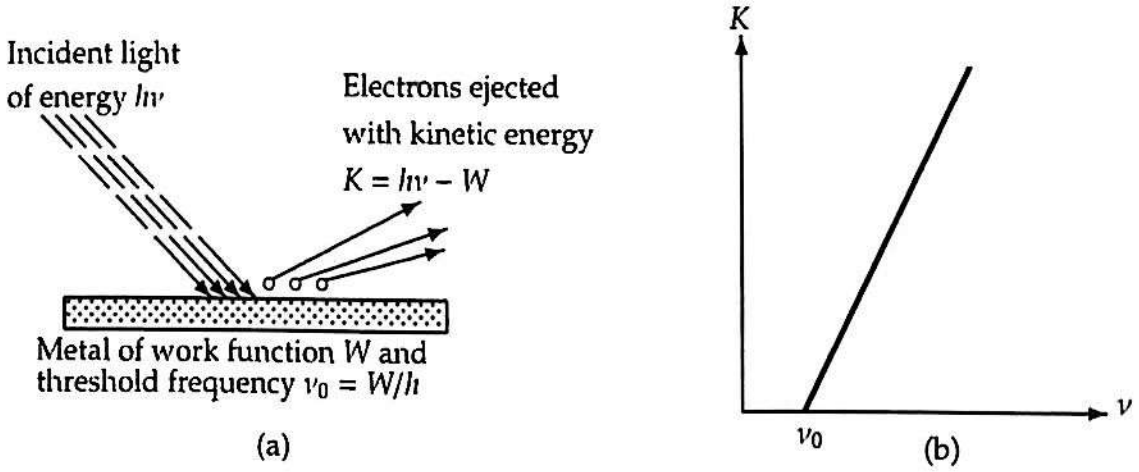


Figure 1.3: (a) Photoelectric effect: when a metal is irradiated with light, electrons may get emitted. (b) Kinetic energy K of the electron leaving the metal when irradiated with a light of frequency ν; when $\nu < \nu_0$ no electron is ejected from the metal regardless of the intensity of the radiation.

That is, since the intensity of an electromagnetic wave is proportional to the square of its amplitude, *any frequency* with sufficient intensity can supply the necessary energy to free the electron from the metal.

But what would happen when using a *weak* light source? According to classical physics, an electron would keep on absorbing energy—at a *continuous rate*—until it gained a sufficient amount; then it would leave the metal. If this argument is to hold, then when using very weak radiation, the photoelectric effect would not take place for a long time, possibly hours, until an electron gradually accumulated the necessary amount of energy. This conclusion, however, disagrees utterly with experimental observation. Experiments were conducted with a light source that was so weak it would have taken several hours for an electron to accumulate the energy needed for its ejection, and yet some electrons were observed to leave the metal *instantly*. Further experiments showed that an increase in intensity (brightness) alone can in no way dislodge electrons from the metal. But by increasing the frequency of the incident radiation beyond a certain threshold, even at very weak intensity, the emission of electrons starts immediately. These experimental facts indicate that the concept of gradual accumulation, or continuous absorption, of energy by the electron, as predicated by classical physics, is indeed erroneous.

Inspired by Planck's quantization of electromagnetic radiation, Einstein succeeded in 1905 in giving a theoretical explanation for the dependence of photoelectric emission on the frequency of the incident radiation. He assumed that light is made of corpuscles each carrying an energy $h\nu$, called *photons*. When a beam of light of frequency ν is incident on a metal, each photon transmits all its energy $h\nu$ to an electron near the surface; in the process, the photon is entirely absorbed by the electron. The electron will thus absorb energy *only* in quanta of energy $h\nu$, irrespective of the intensity of the incident radiation. If $h\nu$ is larger than the metal's *work function* W—the energy required to dislodge the electron from the metal (every metal has free electrons that move from one atom to another; the minimum energy required to free the electron from the metal is called the work function of that metal)—the electron will then be

knocked out of the metal. Hence no electron can be emitted from the metal's surface unless $h\nu > W$:

$$\boxed{h\nu = W + K,} \tag{1.26}$$

where K represents the kinetic energy of the electron leaving the material.

Equation (1.26), which was derived by Einstein, gives the proper explanation to the experimental observation that the kinetic energy of the ejected electron *increases linearly* with the incident frequency ν, as shown in Figure 1.3b:

$$K = h\nu - W = h(\nu - \nu_0), \tag{1.27}$$

where $\nu_0 = W/h$ is called the threshold or cutoff frequency of the metal. Moreover, this relation shows clearly why no electron can be ejected from the metal unless $\nu > \nu_0$: since the kinetic energy cannot be negative, the photoelectric effect cannot occur when $\nu < \nu_0$ regardless of the intensity of the radiation. The ejected electrons acquire their kinetic energy from the excess energy $h(\nu - \nu_0)$ supplied by the incident radiation.

The kinetic energy of the emitted electrons can be experimentally determined as follows. The setup, which was devised by Lenard, consists of the photoelectric metal (cathode) that is placed next to an anode inside an evacuated glass tube. When light strikes the cathode's surface, the electrons ejected will be attracted to the anode, thereby generating a photoelectric current. It was found that *the magnitude of the photoelectric current thus generated is proportional to the intensity of the incident radiation, yet the speed of the electrons does not depend on the radiation's intensity, but on its frequency.* To measure the kinetic energy of the electrons, we simply need to use a varying voltage source and reverse the terminals. When the potential V across the tube is reversed, the liberated electrons will be prevented from reaching the anode; only those electrons with kinetic energy larger than $e|V|$ will make it to the negative plate and contribute to the current. We vary V until it reaches a value V_s, called the *stopping potential*, at which all of the electrons, even the most energetic ones, will be turned back before reaching the collector; hence the flow of photoelectric current ceases completely. The stopping potential V_s is connected to the electrons' kinetic energy by $e|V_s| = \frac{1}{2}m_e v^2 = K$ (in what follows, V_s will implicitly denote $|V_s|$). Thus, the relation (1.27) becomes $eV_s = h\nu - W$ or

$$\boxed{V_s = \frac{h}{e}\nu - \frac{W}{e} = \frac{hc}{e\lambda} - \frac{W}{e}.} \tag{1.28}$$

As shown in Figure 1.4, the shape of the plot of V_s against frequency is a straight line, much like Figure 1.3b with the slope now given by h/e. This shows that the stopping potential depends linearly on the frequency of the incident radiation.

It was Millikan who, in 1916, gave a systematic experimental confirmation to Einstein's photoelectric theory. He produced an extensive collection of photoelectric data using various metals. He verified that Einstein's relation (1.28) reproduced his data exactly. In addition, Millikan found that his empirical value for h, which he obtained by measuring the slope h/e of (1.28) (Figure 1.3b), is equal to Planck's constant to within a 0.5% experimental error.

In summary, the photoelectric effect does provide compelling evidence for the corpuscular nature of the electromagnetic radiation.

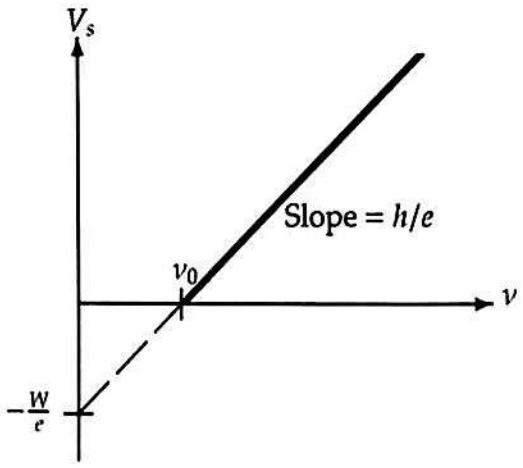


Figure 1.4: Stopping potential V_s as a function of the radiation's frequency ν.

Example 1.2 (Estimation of the Planck constant)
When two ultraviolet beams of wavelengths $\lambda_1 = 80\,\text{nm}$ and $\lambda_2 = 110\,\text{nm}$ fall on a lead surface, they produce photoelectrons with maximum energies 11.390 eV and 7.154 eV, respectively.

(a) Estimate the numerical value of the Planck constant.

(b) Calculate the work function, the cutoff frequency, and the cutoff wavelength of lead.

Solution

(a) From (1.27) we can write the kinetic energies of the emitted electrons as $K_1 = hc/\lambda_1 - W$ and $K_2 = hc/\lambda_2 - W$; the difference between these two expressions is given by $K_1 - K_2 = hc(\lambda_2 - \lambda_1)/(\lambda_1\lambda_2)$ and hence

$$h = \frac{K_1 - K_2}{c}\frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1}. \tag{1.29}$$

Since $1\,\text{eV} = 1.6 \times 10^{-19}\,\text{J}$, the numerical value of h follows at once:

$$h = \frac{(11.390 - 7.154) \times 1.6 \times 10^{-19}\,\text{J}}{3 \times 10^8\,\text{m s}^{-1}} \times \frac{(80 \times 10^{-9}\,\text{m})(110 \times 10^{-9}\,\text{m})}{110 \times 10^{-9}\,\text{m} - 80 \times 10^{-9}\,\text{m}} \simeq 6.627 \times 10^{-34}\,\text{J s}. \tag{1.30}$$

This is a very accurate result indeed.

(b) The work function of the metal can be obtained from either one of the two data

$$\begin{aligned} W = \frac{hc}{\lambda_1} - K_1 &= \frac{6.627 \times 10^{-34}\,\text{J s} \times 3 \times 10^8\,\text{m s}^{-1}}{80 \times 10^{-9}\,\text{m}} - 11.390 \times 1.6 \times 10^{-19}\,\text{J} \\ &= 6.627 \times 10^{-19}\,\text{J} = 4.14\,\text{eV}. \end{aligned} \tag{1.31}$$

The cutoff frequency and wavelength of lead are

$$\nu_0 = \frac{W}{h} = \frac{6.627 \times 10^{-19}\,\text{J}}{6.627 \times 10^{-34}\,\text{J s}} = 10^{15}\,\text{Hz}, \qquad \lambda_0 = \frac{c}{\nu_0} = \frac{3 \times 10^8\,\text{m/s}}{10^{15}\,\text{Hz}} = 300\ \text{nm}. \tag{1.32}$$

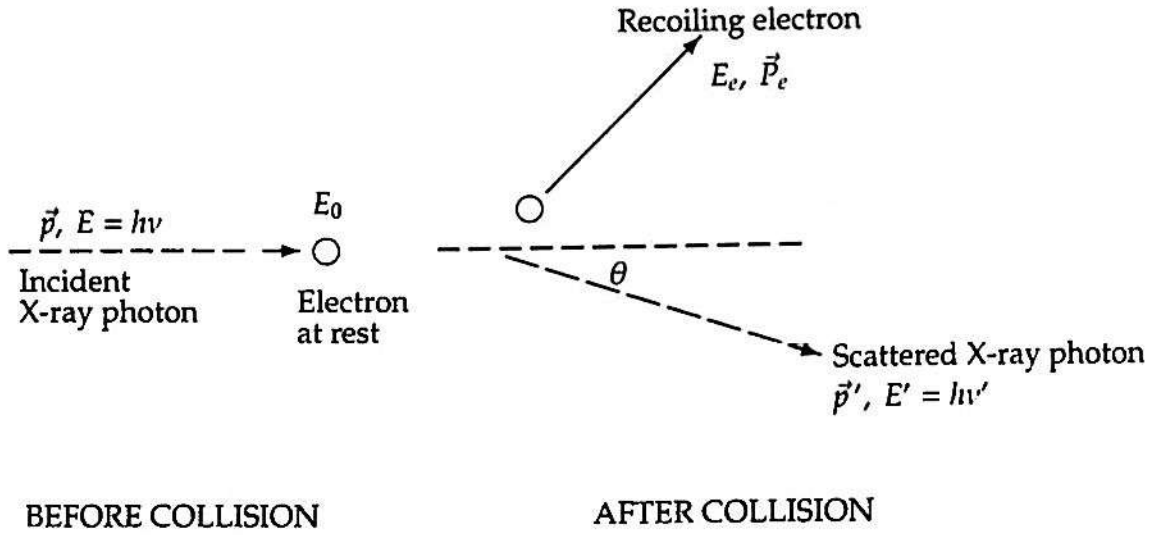


Figure 1.5: Compton scattering of an X-ray photon (of energy $h\nu$ and momentum $\vec{p}$) off a free, stationary electron. After collision, the X-ray photon is scattered at angle θ with energy $h\nu'$.

1.2.3 Compton Effect

In his 1923 experiment, Compton provided the most conclusive confirmation of the particle aspect of radiation. By scattering X-rays off free electrons, he found that the wavelength of the scattered radiation is larger than the wavelength of the incident radiation. This can be explained only by assuming that the X-ray photons behave like particles.

At issue here is to study how X-rays scatter off free electrons. According to classical physics, the incident and scattered radiation should have the same wavelength. This can be viewed as follows. Classically, since the energy of the X-ray radiation is too high to be absorbed by a free electron, the incident X-ray would then provide an oscillatory electric field which sets the electron into oscillatory motion, hence making it radiate light with the same wavelength but with an intensity I that depends on the intensity of the incident radiation I_0 (i.e., $I \propto I_0$). Neither of these two predictions of classical physics is compatible with experiment. The experimental findings of Compton reveal that the wavelength of the scattered X-radiation *increases* by an amount $\Delta\lambda$, called the wavelength shift, and that $\Delta\lambda$ depends not on the intensity of the incident radiation, but only on the scattering angle.

Compton succeeded in explaining his experimental results only after treating the incident radiation as a stream of particles—photons—colliding *elastically* with individual electrons. In this scattering process, which can be illustrated by the elastic scattering of a photon from a free[47] electron (Figure 1.5), the laws of elastic collisions can be invoked, notably the *conservation* of energy and momentum.

Consider that the incident photon, of energy $E = h\nu$ and momentum $p = E/c = h\nu/c$, collides with an electron that is initially at rest. If the photon scatters with a momentum $\vec{p}'$ at an angle[48] θ while the electron recoils with a momentum $\vec{P}_e$, the conservation of linear

[47]When a metal is irradiated with high energy radiation, and at sufficiently high frequencies—as in the case of X-rays—so that $h\nu$ is much larger than the binding energies of the electrons in the metal, these electrons can be considered as free.

[48]Here θ is the angle between $\vec{p}$ and $\vec{p}'$, the photons' momenta before and after collision.

momentum yields

$$\vec{p} = \vec{P}_e + \vec{p}', \tag{1.33}$$

which leads to

$$\vec{P}_e^2 = (\vec{p} - \vec{p}')^2 = p^2 + p'^2 - 2pp' \cos\theta = \frac{h^2}{c^2}\left(\nu^2 + \nu'^2 - 2\nu\nu' \cos\theta\right). \tag{1.34}$$

Let us now turn to the energy conservation. The energies of the electron before and after the collision are given, respectively, by

$$E_0 = m_e c^2, \tag{1.35}$$

$$E_e = \sqrt{\vec{P}_e^2 c^2 + m_e^2 c^4} = h\sqrt{\nu^2 + \nu'^2 - 2\nu\nu' \cos\theta + \frac{m_e^2 c^4}{h^2}}; \tag{1.36}$$

in deriving this relation, we have used (1.34). Since the energies of the incident and scattered photons are given by $E = h\nu$ and $E' = h\nu'$, respectively, conservation of energy dictates that

$$E + E_0 = E' + E_e \tag{1.37}$$

or

$$h\nu + m_e c^2 = h\nu' + h\sqrt{\nu^2 + \nu'^2 - 2\nu\nu' \cos\theta + \frac{m_e^2 c^4}{h^2}}, \tag{1.38}$$

which in turn leads to

$$\nu - \nu' + \frac{m_e c^2}{h} = \sqrt{\nu^2 + \nu'^2 - 2\nu\nu' \cos\theta + \frac{m_e^2 c^4}{h^2}}. \tag{1.39}$$

Squaring both sides of (1.39) and simplifying, we end up with

$$\frac{1}{\nu'} - \frac{1}{\nu} = \frac{h}{m_e c^2}(1 - \cos\theta) = \frac{2h}{m_e c^2} \sin^2\left(\frac{\theta}{2}\right). \tag{1.40}$$

Hence the wavelength shift is given by

$$\boxed{\Delta\lambda = \lambda' - \lambda = \frac{h}{m_e c}(1 - \cos\theta) = 2\lambda_c \sin^2\left(\frac{\theta}{2}\right),} \tag{1.41}$$

where $\lambda_c = h/(m_e c) = 2.426 \times 10^{-12}$ m is called the *Compton wavelength* of the electron. This relation, which connects the initial and final wavelengths to the scattering angle, confirms Compton's experimental observation: the wavelength shift of the X-rays depends only on the angle at which they are scattered and not on the frequency (or wavelength) of the incident photons.

In summary, the Compton effect confirms that photons behave like particles: they collide with electrons like material particles.

Example 1.3 (Compton effect)
High energy photons (γ-rays) are scattered from electrons initially at rest. Assume the photons are backscatterred and their energies are much larger than the electron's rest-mass energy, $E \gg m_e c^2$.

(a) Calculate the wavelength shift.

(b) Show that the energy of the scattered photons is half the rest mass energy of the electron, regardless of the energy of the incident photons.

(c) Calculate the electron's recoil kinetic energy if the energy of the incident photons is 150 MeV.

Solution

(a) In the case where the photons backscatter (i.e., $\theta = \pi$), the wavelength shift (1.41) becomes

$$\Delta\lambda = \lambda' - \lambda = 2\lambda_c \sin^2\left(\frac{\pi}{2}\right) = 2\lambda_c = 4.86 \times 10^{-12}\ \text{m}, \tag{1.42}$$

since $\lambda_c = h/(m_e c) = 2.426 \times 10^{-12}$ m.

(b) Since the energy of the scattered photons E' is related to the wavelength λ' by $E' = hc/\lambda'$, equation (1.42) yields

$$E' = \frac{hc}{\lambda'} = \frac{hc}{\lambda + 2h/(m_e c)} = \frac{m_e c^2}{m_e c^2 \lambda/(hc) + 2} = \frac{m_e c^2}{m_e c^2/E + 2}, \tag{1.43}$$

where $E = hc/\lambda$ is the energy of the incident photons. If $E \gg m_e c^2$ we can approximate (1.43) by

$$E' = \frac{m_e c^2}{2}\left[1 + \frac{m_e c^2}{2E}\right]^{-1} \simeq \frac{m_e c^2}{2} - \frac{(m_e c^2)^2}{4E} \simeq \frac{m_e c^2}{2} = 0.25\ \text{MeV}. \tag{1.44}$$

(c) If $E = 150$ MeV, the kinetic energy of the recoiling electrons can be obtained from conservation of energy

$$K_e = E - E' \simeq 150\ \text{MeV} - 0.25\ \text{MeV} = 149.75\ \text{MeV}. \tag{1.45}$$

1.2.4 Pair Production

We deal here with another physical process which confirms that radiation (the photon) has corpuscular properties.

The theory of quantum mechanics that Schrödinger and Heisenberg proposed works only for nonrelativistic phenomena. This theory, which is called nonrelativistic quantum mechanics, was immensely successful in explaining a wide range of such phenomena. Combining the theory of special relativity with quantum mechanics, Dirac succeeded (1928) in extending quantum mechanics to the realm of relativistic phenomena. The new theory, called relativistic quantum mechanics, predicted the existence of a new particle, the *positron*. This particle, defined as the *antiparticle* of the electron, was predicted to have the same mass as the electron and an equal but opposite (positive) charge.

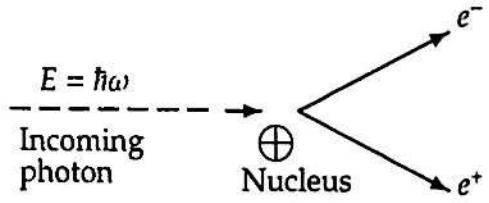


Figure 1.6: Pair production: a highly energetic photon, interacting with a nucleus, disappears and produces an electron and a positron.

Four years after its prediction by Dirac's relativistic quantum mechanics, the positron was discovered by Anderson in 1932 while studying the trails left by cosmic rays in a cloud chamber. When high-frequency electromagnetic radiation passes through a foil, individual photons of this radiation disappear by producing a pair of particles consisting of an electron, e^-, and a positron, e^+: photon$\rightarrow e^- + e^+$. This process is called *pair production*; Anderson obtained such a process by exposing a lead foil to cosmic rays from outer space which contained highly energetic X-rays. It is useless to attempt to explain the pair production phenomenon by means of classical physics, because even nonrelativistic quantum mechanics fails utterly to account for it.

Due to charge, momentum, and energy conservation, pair production cannot occur in empty space. For the process photon$\rightarrow e^- + e^+$ to occur, the photon must interact with an external field such as the Coulomb field of an atomic nucleus to absorb some of its momentum. In the reaction depicted in Figure 1.6, an electron–positron pair is produced when the photon comes near (interacts with) a nucleus at rest; energy conservation dictates that

$$\begin{aligned} \hbar\omega &= E_{e^-} + E_{e^+} + E_N = \left(m_e c^2 + k_{e^-}\right) + \left(m_e c^2 + k_{e^+}\right) + K_N \\ &\simeq 2m_e c^2 + k_{e^-} + k_{e^+}, \end{aligned} \tag{1.46}$$

where $\hbar\omega$ is the energy of the incident photon, $2m_e c^2$ is the sum of the rest masses of the electron and positron, and k_{e^-} and k_{e^+} are the kinetic energies of the electron and positron, respectively. As for $E_N = K_N$, it represents the recoil energy of the nucleus which is purely kinetic. Since the nucleus is very massive compared to the electron and the positron, K_N can be neglected to a good approximation. Note that the photon cannot produce an electron or a positron alone, for electric charge would not be conserved. Also, a massive object, such as the nucleus, must participate in the process to take away some of the photon's momentum.

The inverse of pair production, called pair annihilation, also occurs. For instance, when an electron and a positron collide, they *annihilate* each other and give rise to electromagnetic radiation[49]: $e^- + e^+ \rightarrow$photon. This process explains why positrons do not last long in nature. When a positron is generated in a pair production process, its passage through matter will make it lose some of its energy and it eventually gets annihilated after colliding with an electron. The collision of a positron with an electron produces a hydrogen-like atom, called *positronium*, with a mean lifetime of about 10^{-10} s; positronium is like the hydrogen atom where the proton is replaced by the positron. Note that, unlike pair production, energy

[49]When an electron–positron pair annihilate, they produce at least two photons each having an energy $m_e c^2 = 0.511$ MeV.

and momentum can simultaneously be conserved in pair annihilation processes without any additional (external) field or mass such as the nucleus.

The pair production process is a direct consequence of the mass–energy equation of Einstein $E = mc^2$, which states that pure energy can be converted into mass and vice versa. Conversely, pair annihilation occurs as a result of mass being converted into pure energy. All subatomic particles also have antiparticles (e.g., antiproton). Even neutral particles have antiparticles; for instance, the antineutron is the neutron's antiparticle. Although this text deals only with nonrelativistic quantum mechanics, we have included pair production and pair annihilation, which are relativistic processes, merely to illustrate how radiation interacts with matter, and also to underscore the fact that the quantum theory of Schrödinger and Heisenberg is limited to nonrelativistic phenomena only.

Example 1.4 (Minimum energy for pair production)
Calculate the minimum energy of a photon so that it converts into an electron–positron pair. Find the photon's frequency and wavelength.

Solution
The minimum energy E_{min} of a photon required to produce an electron–positron pair must be equal to the sum of rest mass energies of the electron and positron; this corresponds to the case where the kinetic energies of the electron and positron are zero. Equation (1.46) yields

$$E_{min} = 2m_e c^2 = 2 \times 0.511 \text{ MeV} = 1.02 \text{ MeV}. \tag{1.47}$$

If the photon's energy is smaller than 1.02 MeV, no pair will be produced. The photon's frequency and wavelength can be obtained at once from $E_{min} = h\nu = 2m_e c^2$ and $\lambda = c/\nu$:

$$\nu = \frac{2m_e c^2}{h} = \frac{2 \times 9.1 \times 10^{-31} \text{ kg} \times (3 \times 10^8 \text{m s}^{-1})^2}{6.63 \times 10^{-34} \text{ J s}} = 2.47 \times 10^{20} \text{ Hz}, \tag{1.48}$$

$$\lambda = \frac{c}{\nu} = \frac{3 \times 10^8 \text{m s}^{-1}}{2.47 \times 10^{20} \text{ Hz}} = 1.2 \times 10^{-12} \text{ m}. \tag{1.49}$$

1.3 Wave Aspect of Particles

1.3.1 de Broglie's Hypothesis: Matter Waves

As discussed above—in the photoelectric effect, the Compton effect, and the pair production effect—radiation exhibits particle-like characteristics in addition to its wave nature. In 1923 de Broglie took things even further by suggesting that this wave–particle duality is not restricted to radiation, but must be universal: *all material particles should also display a dual wave–particle behavior*. That is, the wave–particle duality present in light must also occur in matter.

So, starting from the momentum of a photon $p = h\nu/c = h/\lambda$, we can generalize this relation to any material particle[50] with nonzero rest mass: *each material particle of momentum* $\vec{p}$

[50] In classical physics a particle is characterized by its energy E and its momentum $\vec{p}$, whereas a wave is characterized by its wavelength λ and its wave vector $\vec{k} = (2\pi/\lambda)\hat{n}$, where $\hat{n}$ is a unit vector that specifies the direction of propagation of the wave.

behaves as a group of waves (matter waves) whose wavelength λ and wave vector $\vec{k}$ are governed by the speed and mass of the particle

$$\boxed{\lambda = \frac{h}{p} = \frac{h}{mv}, \qquad \vec{k} = \frac{\vec{p}}{\hbar},} \tag{1.50}$$

where $\hbar = h/2\pi$. The expression (1.50), known as the *de Broglie relation*, connects the momentum of a particle with the wavelength and wave vector of the wave corresponding to this particle.

1.3.2 Experimental Confirmation of de Broglie's Hypothesis

de Broglie's idea was confirmed experimentally in 1927 by Davisson and Germer, and later by Thomson, who obtained *interference patterns* with electrons.

1.3.2.1 Davisson–Germer Experiment

In their experiment, Davisson and Germer scattered a 54 eV monoenergetic beam of electrons from a nickel (Ni) crystal. The electron source and detector were symmetrically located with respect to the crystal's normal, as indicated in Figure 1.7; this is similar to the Bragg setup for X-ray diffraction by a grating. What Davisson and Germer found was that, although the electrons are scattered in all directions from the crystal, the intensity was a minimum at $\theta = 35°$ and a maximum at $\theta = 50°$; that is, the bulk of the electrons scatter only in well-specified directions. They showed that the pattern persisted even when the intensity of the beam was so low that the incident electrons were sent one at a time. This can only result from a constructive interference of the scattered electrons. So, instead of the diffuse distribution pattern that results from material particles, the reflected electrons formed diffraction patterns that were identical with Bragg's X-ray diffraction by a *grating*. In fact, the intensity maximum of the scattered electrons in the Davisson–Germer experiment corresponds to the first maximum ($n = 1$) of the Bragg formula,

$$n\lambda = 2d \sin \phi, \tag{1.51}$$

where d is the spacing between the Bragg planes, ϕ is the angle between the incident ray and the crystal's reflecting planes, θ is the angle between the incident and scattered beams (d is given in terms of the separation D between successive atomic layers in the crystal by $d = D \sin \theta$).

For an Ni crystal, we have $d = 0.091$ nm, since $D = 0.215$ nm. Since only one maximum is seen at $\theta = 50°$ for a mono-energetic beam of electrons of kinetic energy 54 eV, and since $2\phi + \theta = \pi$ and hence $\sin \phi = \cos(\theta/2)$ (Figure 1.7), we can obtain from (1.51) the wavelength associated with the scattered electrons:

$$\lambda = \frac{2d}{n} \sin \phi = \frac{2d}{n} \cos \frac{1}{2}\theta = \frac{2 \times 0.091 \text{ nm}}{1} \cos 25° = 0.165 \text{ nm}. \tag{1.52}$$

Now, let us look for the numerical value of λ that results from de Broglie's relation. Since the kinetic energy of the electrons is $K = 54$ eV, and since the momentum is $p = \sqrt{2m_e K}$ with

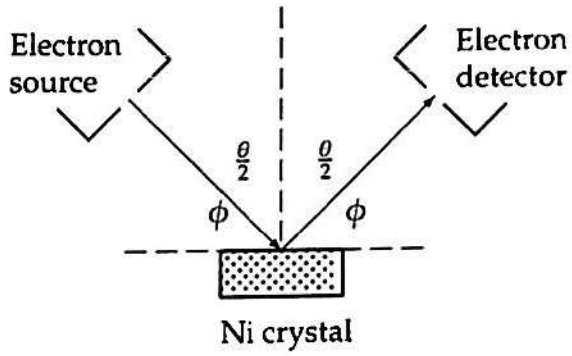


Figure 1.7: Davisson–Germer experiment: electrons strike the crystal's surface at an angle ϕ; the detector, symmetrically located from the electron source, measures the number of electrons scattered at an angle θ, where θ is the angle between the incident and scattered electron beams.

$m_e c^2 = 0.511$ MeV (the rest mass energy of the electron) and $\hbar c \simeq 197.33$ eV nm, we can show that the de Broglie wavelength is

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2m_e K}} = \frac{2\pi\hbar c}{\sqrt{2m_e c^2 K}} = 0.167 \text{ nm}, \tag{1.53}$$

which is in excellent agreement with the experimental value (1.52).

We have seen that the scattered electrons in the Davisson–Germer experiment produced interference fringes that were identical to those of Bragg's X-ray diffraction. Since the Bragg formula provided an accurate prediction of the electrons' interference fringes, the motion of an electron of momentum $\vec{p}$ must be described by means of a plane wave

$$\psi(\vec{r}, t) = Ae^{i(\vec{k}\cdot\vec{r}-\omega t)} = Ae^{i(\vec{p}\cdot\vec{r}-Et)/\hbar}, \tag{1.54}$$

where A is a constant, $\vec{k}$ is the wave vector of the plane wave, and ω is its angular frequency; the wave's parameters, $\vec{k}$ and ω, are related to the electron's momentum $\vec{p}$ and energy E by means of de Broglie's relations: $\vec{k} = \vec{p}/\hbar$, $\omega = E/\hbar$.

We should note that, inspired by de Broglie's hypothesis, Schrödinger constructed the theory of wave mechanics which deals with the dynamics of microscopic particles. *He described the motion of particles by means of a wave function* $\psi(\vec{r}, t)$ which corresponds to the de Broglie wave of the particle. We will deal with the physical interpretation of $\psi(\vec{r}, t)$ in the following section.

1.3.2.2 Thomson Experiment

In the Thomson experiment (Figure 1.8), electrons were diffracted through a polycrystalline thin film. Diffraction fringes were also observed. This result confirmed again the wave behavior of electrons.

The Davisson–Germer experiment has inspired others to obtain diffraction patterns with a large variety of particles. Interference patterns were obtained with bigger and bigger particles such as neutrons, protons, helium atoms, and hydrogen molecules. de Broglie wave

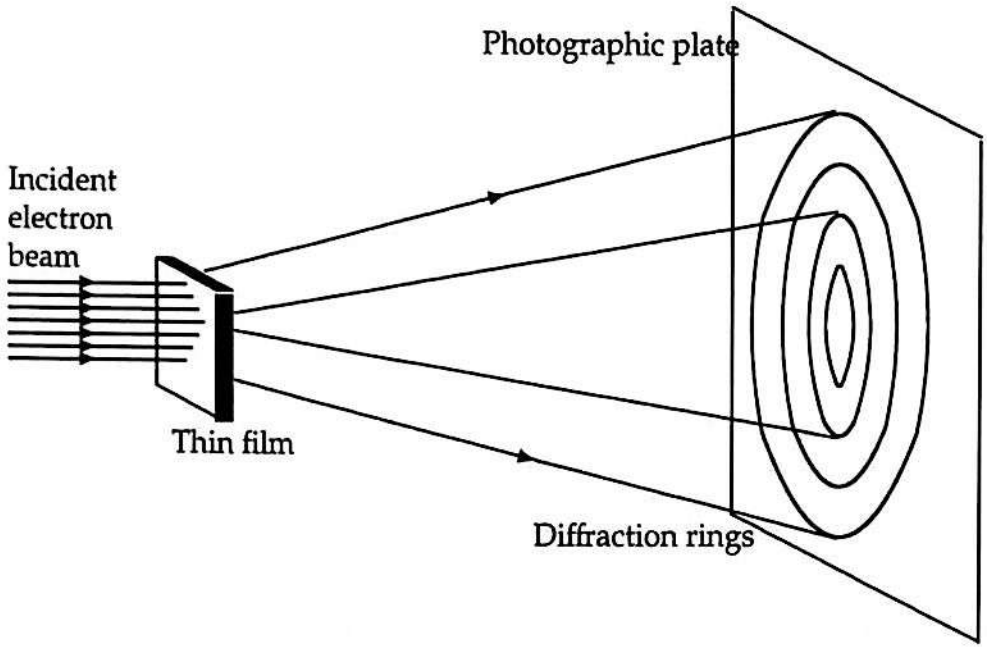


Figure 1.8: Thomson experiment: diffraction of electrons through a thin film of polycrystalline material yields fringes that usually result from light diffraction.

interference of carbon 60 (C60) molecules were recently[51] observed by diffraction at a material absorption grating; these observations supported the view that each C60 molecule interferes only with itself (a C60 molecule is nearly a classical object).

1.3.3 Matter Waves for Macroscopic Objects

We have seen that microscopic particles, such as electrons, display wave behavior. What about macroscopic objects? Do they also display wave features? They surely do. Although macroscopic material particles display wave properties, the corresponding wavelengths are too small to detect; being very massive[52], macroscopic objects have extremely small wavelengths. At the microscopic level, however, the waves associated with material particles are of the same size or exceed the size of the system. Microscopic particles therefore exhibit clearly discernible wave-like aspects.

The general rule is: whenever the de Broglie wavelength of an object is in the range of, or exceeds, its size, the wave nature of the object is detectable and hence cannot be neglected. But if its de Broglie wavelength is much too small compared to its size, the wave behavior of this object is undetectable. For a quantitative illustration of this general rule, let us calculate in the following example the wavelengths corresponding to two particles, one microscopic and the other macroscopic.

Example 1.5 (Matter waves for microscopic and macroscopic systems)
Calculate the de Broglie wavelength for

[51] Markus Arndt, *et al.*, "Wave–Particle Duality of C60 Molecules", *Nature*, V401, n6754, 680 (Oct. 14, 1999).

[52] Very massive compared to microscopic particles. For instance, the ratio between the mass of an electron and a 100 g bullet is infinitesimal: $m_e/m_b \simeq 10^{-29}$.

(a) a proton of kinetic energy 70 MeV kinetic energy and
(b) a 100 g bullet moving at 900 $\mathrm{m\,s^{-1}}$.

Solution

(a) Since the kinetic energy of the proton is $T = p^2/(2m_p)$, its momentum is $p = \sqrt{2Tm_p}$. The de Broglie wavelength is $\lambda_p = h/p = h/\sqrt{2Tm_p}$. To calculate this quantity numerically, it is more efficient to introduce the well-known quantity $\hbar c \simeq 197$ MeV fm and the rest mass of the proton $m_pc^2 = 938.3$ MeV, where c is the speed of light:

$$\lambda_p = 2\pi\frac{\hbar c}{pc} = 2\pi\frac{\hbar c}{\sqrt{2Tm_pc^2}} = 2\pi\frac{197\ \text{MeV fm}}{\sqrt{2\times 938.3\times 70\ \text{MeV}^2}} = 3.4\times 10^{-15}\ \text{m}. \tag{1.55}$$

(b) As for the bullet, its de Broglie wavelength is $\lambda_b = h/p = h/(mv)$ and since $h = 6.626\times 10^{-34}$ J s, we have

$$\lambda_b = \frac{h}{mv} = \frac{6.626\times 10^{-34}\ \text{J s}}{0.1\ \text{kg}\times 900\ \text{m s}^{-1}} = 7.4\times 10^{-36}\ \text{m}. \tag{1.56}$$

The ratio of the two wavelengths is $\lambda_b/\lambda_p \simeq 2.2\times 10^{-21}$. Clearly, the wave aspect of this bullet lies beyond human observational abilities. As for the wave aspect of the proton, it cannot be neglected; its de Broglie wavelength of 3.4×10^{-15} m has the same order of magnitude as the size of a typical atomic nucleus.

We may conclude that, whereas the wavelengths associated with *microscopic* systems are *finite* and display easily detectable wave-like patterns, the wavelengths associated with *macroscopic* systems are *infinitesimally small* and display no discernible wave-like behavior. So, when the wavelength approaches zero, the wave-like properties of the system disappear. In such cases of infinitesimally small wavelengths, *geometrical optics* should be used to describe the motion of the object, for the wave associated with it behaves as a *ray*.

1.4 Particles versus Waves

In this section we are going to study the properties of particles and waves within the contexts of classical and quantum physics. The experimental setup to study these aspects is the *double-slit experiment*, which consists of a source S (S can be a source of material particles or of waves), a wall with two slits S_1 and S_2, and a back screen equipped with counters that record whatever arrives at it from the slits.

1.4.1 Classical View of Particles and Waves

In classical physics, particles and waves are mutually exclusive; they exhibit completely different behaviors. While the full description of a particle requires only one parameter, the position vector $\vec{r}(t)$, the complete description of a wave requires two, the amplitude and the phase. For instance, three-dimensional plane waves can be described by wave functions $\psi(\vec{r},t)$:

$$\psi(\vec{r},t) = Ae^{i(\vec{k}\cdot\vec{r}-\omega t)} = Ae^{i\phi}, \tag{1.57}$$

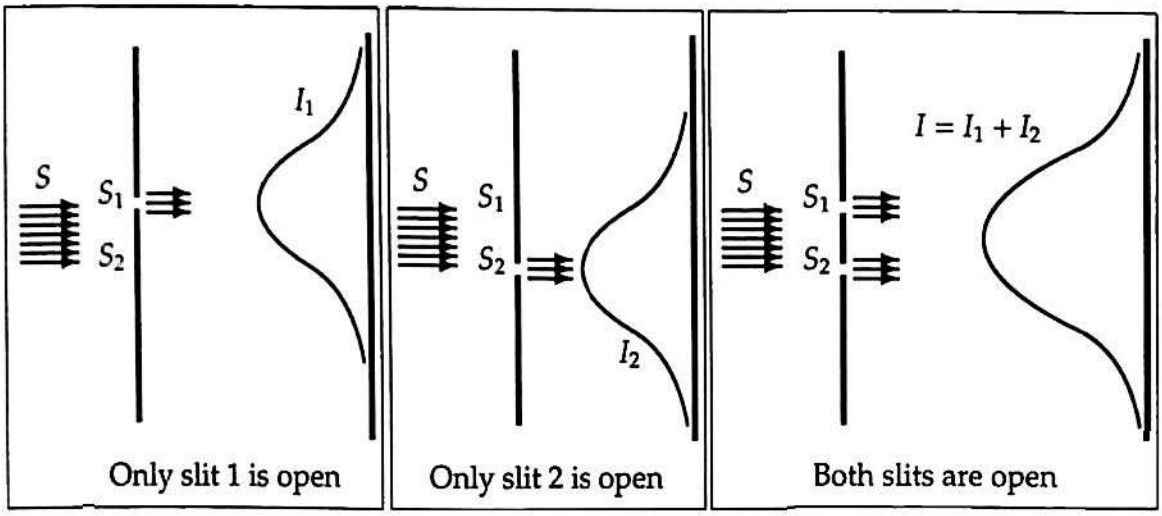


Figure 1.9: The double-slit experiment with *particles*: S is a source of *bullets*; I_1 and I_2 are the intensities recorded on the screen, respectively, when only S_1 is open and then when only S_2 is open. When both slits are open, the total intensity is $I = I_1 + I_2$.

where A is the amplitude of the wave and ϕ is its phase ($\vec{k}$ is the wave vector and ω is the angular frequency). We may recall the physical meaning of ψ: the intensity of the wave is given by $I = |\psi|^2$.

(a) S is a source of streams of bullets

Consider three different experiments as displayed in Figure 1.9, in which a source S fires a stream of bullets; the bullets are assumed to be indestructible and hence arrive on the screen in identical lumps. In the first experiment, only slit S_1 is open; let $I_1(y)$ be the corresponding intensity collected on the screen (the number of bullets arriving per second at a given point y). In the second experiment, let $I_2(y)$ be the intensity collected on the screen when only S_2 is open. In the third experiments, if S_1 and S_2 are both open, the total intensity collected on the screen behind the two slits must be equal to the sum of I_1 and I_2:

$$I(y) = I_1(y) + I_2(y). \tag{1.58}$$

(b) S is a source of waves

Now, as depicted in Figure 1.10, S is a source of waves (e.g., light or water waves). Let I_1 be the intensity collected on the screen when only S_1 is open and I_2 be the intensity when only S_2 is open. Recall that a wave is represented by a complex function ψ, and its intensity is proportional to its amplitude (e.g., height of water or electric field) squared: $I_1 = |\psi_1|^2$, $I_2 = |\psi_2|^2$. When both slits are open, the total intensity collected on the screen displays an *interference* pattern; hence it cannot be equal to the sum of I_1 and I_2. The amplitudes, not the intensities, must add: the total amplitude ψ is the sum of ψ_1 and ψ_2; hence the total intensity is given by

$$\begin{aligned} I = |\psi_1 + \psi_2|^2 &= |\psi_1|^2 + |\psi_2|^2 + \left(\psi_1^*\psi_2 + \psi_2^*\psi_1\right) = I_1 + I_2 + 2\mathrm{Re}(\psi_1^*\psi_2) \\ &= I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\delta, \end{aligned} \tag{1.59}$$

where δ is the phase difference between ψ_1 and ψ_2, and $2\sqrt{I_1 I_2}\cos\delta$ is an oscillating term,

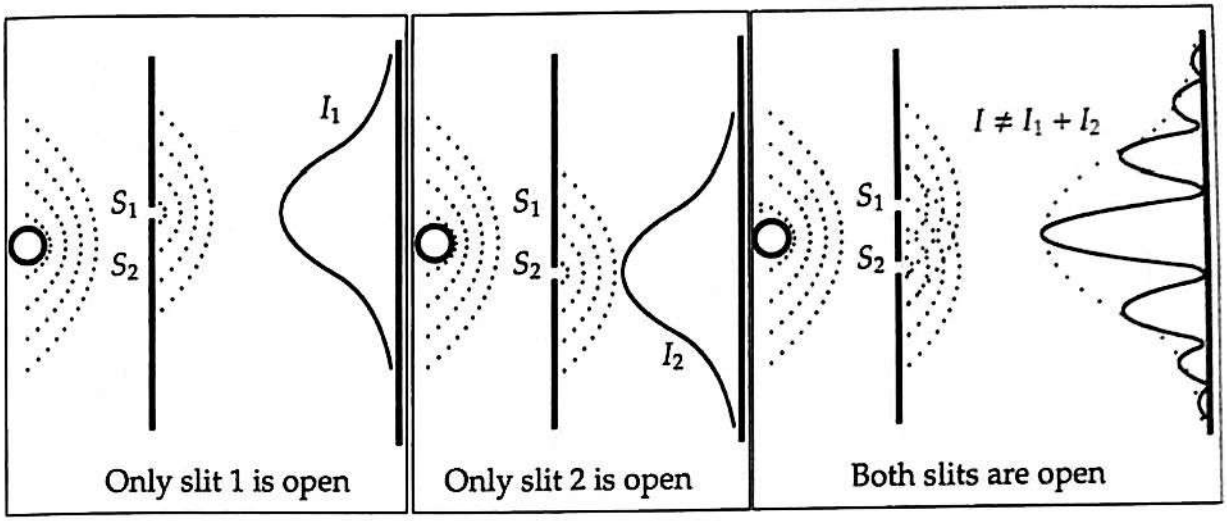


Figure 1.10: The double-slit experiment: S is a source of *waves*, I_1 and I_2 are the intensities recorded on the screen when only S_1 is open, and then when only S_2 is open, respectively. When both slits are open, the total intensity is no longer equal to the sum of I_1 and I_2; an *oscillating* term has to be added.

which is responsible for the interference pattern (Figure 1.10). So the resulting intensity distribution cannot be predicted from I_1 or from I_2 alone, for it depends on the phase δ, which cannot be measured when only one slit is open (δ can be calculated from the slits separation or from the observed intensities I_1, I_2 and I).

Conclusion: Classically, waves exhibit interference patterns, particles do not. When two noninteracting streams of particles combine in the same region of space, their intensities add; when waves combine, their amplitudes add but their intensities do not.

1.4.2 Quantum View of Particles and Waves

Let us now discuss the double-slit experiment with quantum material particles such as electrons. Figure 1.11 shows three different experiments where the source S shoots a stream of electrons, first with only S_1 open, then with only S_2 open, and finally with both slits open. In the first two cases, the distributions of the electrons on the screen are smooth; the sum of these distributions is also smooth, a bell-shaped curve like the one obtained for classical particles (Figure 1.9).

But when both slits are open, we see a rapid variation in the distribution, an *interference pattern*. So in spite of their discreteness, the electrons seem to interfere with themselves; this means that each electron seems to have gone through both slits at once! One might ask, if an electron cannot be split, how can it appear to go through both slits at once? Note that this interference pattern has nothing to do with the intensity of the electron beam. In fact, experiments were carried out with beams so weak that the electrons were sent one at a time (i.e., each electron was sent only after the previous electron has reached the screen). In this case, if both slits were open and if we wait long enough so that sufficient impacts are collected on the screen, the interference pattern appears again.

The crucial question now is to find out the slit through which the electron went. To answer

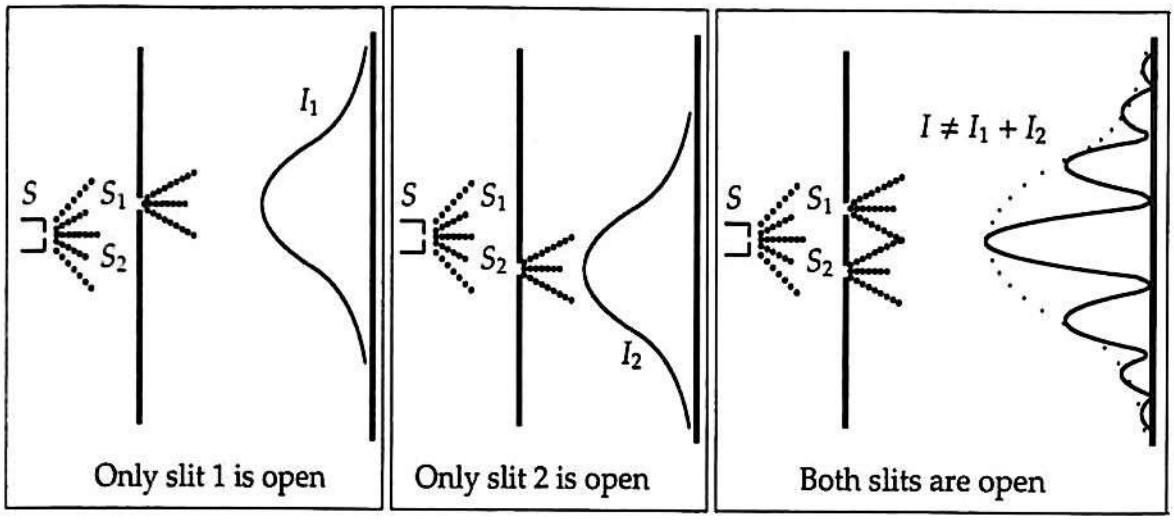


Figure 1.11: The double-slit experiment: S is a source of *electrons*, I_1 and I_2 are the intensities recorded on the screen when only S_1 is open, and then when only S_2 is open, respectively. When both slits are open, the total intensity is equal to the sum of I_1, I_2 and an *oscillating* term.

this query, an experiment can be performed to watch the electrons as they leave the slits. It consists of placing a strong light source behind the wall containing the slits, as shown in Figure 1.12. We place Geiger counters all over the screen so that whenever an electron reaches the screen we hear a click on the counter.

Since electric charges scatter light, whenever an electron passes through either of the slits, on its way to the counter, it will scatter light to our eyes. So, whenever we hear a click on the counter, we see a flash near *either* S_1 *or* S_2 *but never near both at once.* After recording the various counts with both slits open, we find out that the distribution is similar to that of classical bullets in Figure 1.9: the interference pattern has disappeared! But if we turn off the light source, the interference pattern appears again.

From this experiment we conclude that the mere act of looking at the electrons immensely affects their distribution on the screen. Clearly, electrons are very delicate: their motion gets modified when one watches them. This is the very quantum mechanical principle which states that *measurements interfere with the states of microscopic objects.* One might think of turning down the brightness (intensity) of the light source so that it is weak enough not to disturb the electrons. We find that the light scattered from the electrons, as they pass by, does not get weaker; the same sized flash is seen, but only every once in a while. This means that, at low brightness levels, we miss some electrons: we hear the click from the counter but see no flash at all. At still lower brightness levels, we miss most of the electrons. We conclude, in this case, that some electrons went through the slits without being seen, because there were no photons around at the right moment to catch them. This process is important because it confirms that light has particle properties: light also arrives in lumps (photons) at the screen.

Two distribution profiles are compiled from this dim light source experiment, one corresponding to the electrons that were seen and the other to the electrons that were not seen (but heard on the counter). The first distribution contains no interference (i.e., it is similar to classical bullets); but the second distribution displays an interference pattern. This results from the fact that when the electrons are not seen, they display interference. When we do not

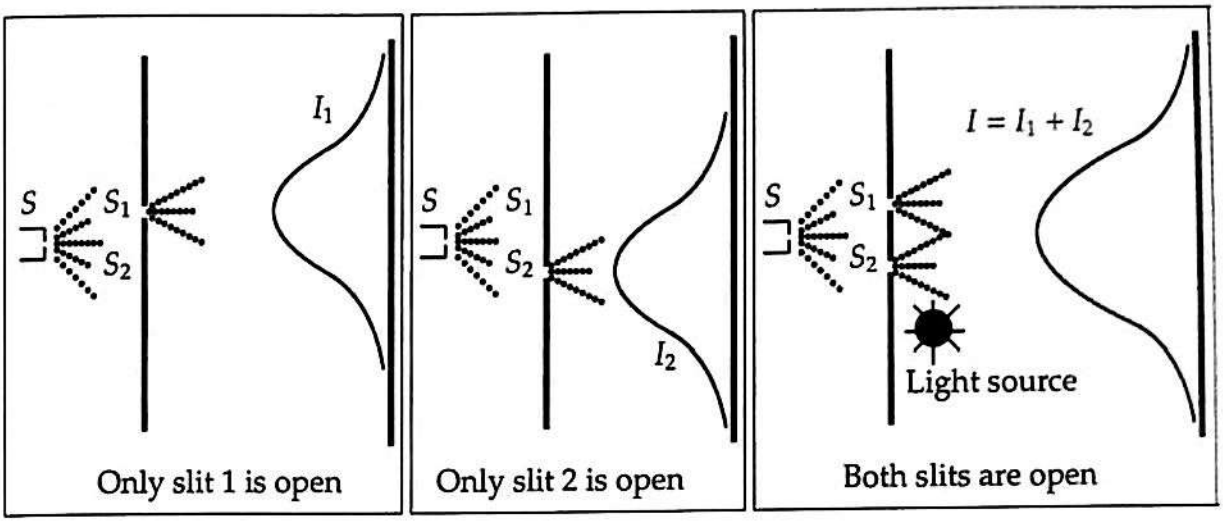


Figure 1.12: The double-slit experiment: S is a source of *electrons*. *A light source* is placed behind the wall containing S_1 and S_2. When both slits are open, the interference pattern is destroyed and the total intensity is $I = I_1 + I_2$.

see the electron, no photon has disturbed it but when we see it, a photon has disturbed it.

For the electrons that display interference, it is impossible to identify the slit that each electron had gone through. This experimental finding introduces a new fundamental concept: the microphysical world is *indeterministic*. Unlike classical physics, where we can follow accurately the particles along their trajectories, we cannot follow a microscopic particle along its motion nor can we determine its path. It is technically impossible to perform such detailed tracing of the particle's motion. Such results inspired Heisenberg to postulate the uncertainty principle, which states that *it is impossible to design an apparatus which allows us to determine the slit that the electron went through without disturbing the electron enough to destroy the interference pattern* (we shall return to this principle later).

The interference pattern obtained from the double-slit experiment indicates that electrons display both particle and wave properties. When electrons are observed or detected one by one, they behave like particles, but when they are detected after many measurements (distribution of the detected electrons), they behave like waves of wavelength $\lambda = h/p$ and display an interference pattern.

1.4.3 Wave–Particle Duality: Complementarity

The various experimental findings discussed so far—blackbody radiation, photoelectric and Compton effect, pair production, Davisson–Germer, Thomson, and the double-slit experiments—reveal that photons, electrons, and any other microscopic particles behave unlike classical particles and unlike classical waves. These findings indicate that, at the microscopic scale, nature can display particle behavior as well as wave behavior. The question now is, how can something behave as a particle and as a wave at the same time? Aren't these notions mutually exclusive? In the realm of classical physics the answer is yes, but not in quantum mechanics. This *dual* behavior can in no way be reconciled within the context of classical physics, for particles and waves are mutually exclusive entities.

The theory of quantum mechanics, however, provides the proper framework for reconciling the particle and wave aspects of matter. By using a wave function $\psi(\vec{r}, t)$ (see (1.54)) to describe material particles such as electrons, quantum mechanics can simultaneously make statements about the particle behavior and the wave behavior of microscopic systems. It combines the quantization of energy or intensity with a wave description of matter. That is, it uses both particle and wave pictures to describe the same material particle.

Our ordinary concepts of particles or waves are thus inadequate when applied to microscopic systems. These two concepts, which preclude each other in the macroscopic realm, do not strictly apply to the microphysical world. No longer valid at the microscopic scale is the notion that a wave cannot behave as a particle and vice versa. The true reality of a quantum system is that it is neither a pure particle nor a pure wave. The particle and wave aspects of a quantum system manifest themselves only when subjected to, or intruded on by, penetrating means of observation (any procedure of penetrating observation would destroy the initial state of the quantum system; for instance, the mere act of looking at an electron will knock it out of its orbit). Depending on the type of equipment used to observe an electron, the electron has the capacity to display either "grain" or wave features. As illustrated by the double-slit experiment, if we wanted to look at the particle aspect of the electron, we would need only to block one slit (or leave both slits open but introduce an observational apparatus), but if we were interested only in its wave features, we would have to leave both slits open and not intrude on it by observational tools. This means that both the "grain" and "wave" features are embedded into the electron, and by modifying the probing tool, we can suppress one aspect of the electron and keep the other. An experiment designed to isolate the particle features of a quantum system gives no information about its wave features, and vice versa. When we subject an electron to Compton scattering, we observe only its particle aspects, but when we involve it in a diffraction experiment (as in Davisson–Germer, Thomson, or the double-slit experiment), we observe its wave behavior only. So if we measure the particle properties of a quantum system, this will destroy its wave properties, and vice versa. Any measurement gives either one property or the other, but never both at once. We can get either the wave property or the particle aspect but not both of them together.

Microscopic systems, therefore, are neither pure particles nor pure waves, they are both. The particle and wave manifestations do not contradict or preclude one another, but, as suggested by Bohr, they are just *complementary*. Both concepts are complementary in describing the true nature of microscopic systems. Being complementary features of microscopic matter, particles and waves are equally important for a complete description of quantum systems. From here comes the essence of the *complementarity principle*.

We have seen that when the *rigid* concept of *either/or* (i.e., either a particle or a wave) is indiscriminately applied or imposed on quantum systems, we get into trouble with reality. Without the complementarity principle, quantum mechanics would not have been in a position to produce the accurate results it does.

Note: In the case of light, although quantum mechanics predicts that light behaves simultaneously as a particle and as a wave, there has never been an experiment that managed to directly capture these two aspects of light at the same time. In the phenomena studied above –the blackbody radiation, the photoelectric effect, Compton effect, the pair production effect, and the double-slit experiment – light was shown experimentally to display either wave or particle aspects always at different times, but never simultaneously. That was the case until recently. In 2015, researchers have managed to devise an experiment that managed to capture

the first ever picture where light behaves both as a wave and as a particle at the same time[53]. In this experiment, where electrons were used to "image" light, a pulse of laser light was fired at a tiny silver nanowire thereby adding energy to the charged particles in the nanowire, causing them to oscillate. Light traveling along this nanowire in two opposite directions gives rise to a new wave – a standing wave – that becomes a source of light radiating around the nanowire. Then a stream of electrons were shot close to the nanowire and used to image the standing wave of light. As the electrons interacted with the standing wave (confined standing wave of light on the nanowire), they either sped up or slowed down; this change of speed of the electrons appears as an exchange of energy packets (quanta) between electrons and photons. Whereas the imaged standing wave acts as a fingerprint of the wave-aspect of light, the occurrence of these energy packets shows that the light on the nanowire behaves as a particle.

1.4.4 Principle of Linear Superposition

How do we account mathematically for the existence of the interference pattern in the double-slit experiment with material particles such as electrons? An answer is offered by the *superposition principle*. The interference results from the superposition of the waves emitted by slits 1 and 2. If the functions $\psi_1(\vec{r},t)$ and $\psi_2(\vec{r},t)$, which denote the waves reaching the screen emitted respectively by slits 1 and 2, represent two physically possible states of the system, then any linear superposition

$$\psi(\vec{r},t) = \alpha_1\psi_1(\vec{r},t) + \alpha_2\psi_2(\vec{r},t) \tag{1.60}$$

also represents a physically possible outcome of the system; α_1 and α_2 are complex constants. This is the superposition principle. The intensity produced on the screen by opening only slit 1 is $\left|\psi_1(\vec{r},t)\right|^2$ and it is $\left|\psi_2(\vec{r},t)\right|^2$ when only slit 2 is open. When both slits are open, the intensity is

$$\begin{aligned}\left|\psi(\vec{r},t)\right|^2 &= \left|\psi_1(\vec{r},t) + \psi_2(\vec{r},t)\right|^2 \\ &= \left|\psi_1(\vec{r},t)\right|^2 + \left|\psi_2(\vec{r},t)\right|^2 + \psi_1^*(\vec{r},t)\psi_2(\vec{r},t) + \psi_1(\vec{r},t)\psi_2^*(\vec{r},t),\end{aligned} \tag{1.61}$$

where the asterisk denotes the complex conjugate. Note that (1.61) is not equal to the sum of $\left|\psi_1(\vec{r},t)\right|^2$ and $\left|\psi_2(\vec{r},t)\right|^2$; it contains an additional term $\psi_1^*(\vec{r},t)\psi_2(\vec{r},t) + \psi_1(\vec{r},t)\psi_2^*(\vec{r},t)$. This is the very term which gives rise in the case of electrons to an interference pattern similar to light waves. The interference pattern therefore results from the existence of a phase shift between $\psi_1(\vec{r},t)$ and $\psi_2(\vec{r},t)$. We can measure this phase shift from the interference pattern, but we can in no way measure the phases of ψ_1 and ψ_2 separately.

We can summarize the double-slit results in three principles:

- Intensities add for classical particles: $I = I_1 + I_2$.
- Amplitudes, not intensities, add for quantum particles: $\psi(\vec{r},t) = \psi_1(\vec{r},t) + \psi_2(\vec{r},t)$; this gives rise to interference.

[53] L. Piazza, T.T.A. Lummen, E. Quiñonez, Y. Murooka, B.W. Reed, B. Barwick, and F. Carbone, *Simultaneous observation of the quantization and the interference pattern of a plasmonic near-field*, *Nature Communications* 6, 02 March 2015. Article number: 6407 (2015).

- Whenever one attempts to determine experimentally the outcome of individual events for microscopic material particles (such as trying to specify the slit through which an electron has gone), the interference pattern gets destroyed. In this case the intensities add in much the same way as for classical particles: $I = I_1 + I_2$.

1.5 Indeterministic Nature of the Microphysical World

Let us first mention two important experimental findings that were outlined above. On the one hand, the Davisson–Germer and the double-slit experiments have shown that microscopic material particles do give rise to interference patterns. To account for the interference pattern, we have seen that it is imperative to describe microscopic particles by means of waves. Waves are not localized in space. As a result, we have to give up on accuracy to describe microscopic particles, for waves give at best a probabilistic account. On the other hand, we have seen in the double-slit experiment that it is impossible to trace the motion of individual electrons; there is no experimental device that would determine the slit through which a given electron has gone. Not being able to predict single events is a stark violation of a founding principle of classical physics: predictability or determinacy. These experimental findings inspired Heisenberg to postulate the indeterministic nature of the microphysical world and Born to introduce the probabilistic interpretation of quantum mechanics.

1.5.1 Heisenberg's Uncertainty Principle

According to classical physics, given the initial conditions and the forces acting on a system, the future behavior (unique path) of this physical system can be determined exactly. That is, if the initial coordinates $\vec{r}_0$, velocity $\vec{v}_0$, and all the forces acting on the particle are known, the position $\vec{r}(t)$ and velocity $\vec{v}(t)$ are uniquely determined by means of Newton's second law. *Classical physics* is thus completely *deterministic*.

Does this deterministic view hold also for the microphysical world? Since a particle is represented within the context of quantum mechanics by means of a wave function corresponding to the particle's wave, and since wave functions cannot be localized, then a microscopic particle is somewhat spread over space and, unlike classical particles, cannot be localized in space. In addition, we have seen in the double-slit experiment that it is impossible to determine the slit that the electron went through without disturbing it. The classical concepts of exact position, exact momentum, and unique path of a particle therefore make no sense at the microscopic scale. This is the essence of Heisenberg's uncertainty principle.

In its original form, Heisenberg's uncertainty principle[54] states that: *If the x-component of the momentum of a particle is measured with an uncertainty Δp_x, then its x-position cannot, at the same time, be measured more accurately than $\Delta x = \hbar/(2\Delta p_x)$.* The three-dimensional form of the uncertainty relations for position and momentum can be written as follows:

$$\boxed{\Delta x \Delta p_x \geq \frac{\hbar}{2}, \qquad \Delta y \Delta p_y \geq \frac{\hbar}{2}, \qquad \Delta z \Delta p_z \geq \frac{\hbar}{2}.} \tag{1.62}$$

This principle indicates that, although it is possible to measure the momentum or position of a particle accurately, it is not possible to measure these two observables *simultaneously* to

[54] W. Heisenberg, *Zeits. für Physik*, Vol. 43, 172 (1927).

an *arbitrary* accuracy. That is, we cannot localize a microscopic particle without giving to it a rather large momentum. We cannot measure the position without disturbing it; there is no way to carry out such a measurement passively as it is bound to change the momentum. To understand this, consider measuring the position of a macroscopic object (e.g., a car) and the position of a microscopic system (e.g., an electron in an atom). On the one hand, to locate the position of a macroscopic object, you need simply to observe it; the light that strikes it and gets reflected to the detector (your eyes or a measuring device) can in no measurable way affect the motion of the object. On the other hand, to measure the position of an electron in an atom, you must use radiation of very short wavelength (the size of the atom). The energy of this radiation is high enough to change tremendously the momentum of the electron; the mere observation of the electron affects its motion so much that it can knock it entirely out of its orbit. It is therefore impossible to determine the position and the momentum simultaneously to arbitrary accuracy. If a particle were localized, its wave function would become zero everywhere else and its wave would then have a very short wavelength. According to de Broglie's relation $p = h/\lambda$, the momentum of this particle will be rather *high*. Formally, this means that if a particle is accurately localized (i.e., $\Delta x \to 0$), there will be total uncertainty about its momentum (i.e., $\Delta p_x \to \infty$). To summarize, since all quantum phenomena are described by waves, we have no choice but to accept limits on our ability to measure simultaneously any two complementary variables.

Heisenberg's uncertainty principle can be generalized to any pair of complementary, or canonically conjugate, dynamical variables: *it is impossible to devise an experiment that can measure simultaneously two complementary variables to arbitrary accuracy* (if this were ever achieved, the theory of quantum mechanics would collapse).

Energy and time, for instance, form a pair of complementary variables. Their simultaneous measurement must obey the time–energy uncertainty relation:

$$\boxed{\Delta E \Delta t \geq \frac{\hbar}{2}.} \tag{1.63}$$

This relation states that if we make two measurements of the energy of a system and if these measurements are separated by a time interval Δt, the measured energies will differ by an amount ΔE which can in no way be smaller than $\hbar/\Delta t$. If the time interval between the two measurements is large, the energy difference will be small. This can be attributed to the fact that, when the first measurement is carried out, the system becomes perturbed and it takes it a long time to return to its initial, unperturbed state. This expression is particularly useful in the study of decay processes, for it specifies the relationship between the mean lifetime and the energy width of the excited states.

We see that, in sharp contrast to classical physics, *quantum mechanics* is a completely *indeterministic* theory. Asking about the position or momentum of an electron, one cannot get a definite answer; only a *probabilistic* answer is possible. According to the uncertainty principle, if the position of a quantum system is well defined, its momentum will be totally undefined. In this context, the uncertainty principle has clearly brought down one of the most sacrosanct concepts of classical physics: the deterministic nature of Newtonian mechanics.

Example 1.6 (Uncertainties for microscopic and macroscopic systems)
Estimate the uncertainty in the position of (a) a neutron moving at $5 \times 10^6 \text{m s}^{-1}$ and (b) a 50 kg person moving at 2m s^{-1}.

Solution

(a) Using (1.62), we can write the position uncertainty as

$$\Delta x \geq \frac{\hbar}{2\Delta p} \simeq \frac{\hbar}{2m_n v} = \frac{1.05 \times 10^{-34}\ \text{J s}}{2 \times 1.65 \times 10^{-27}\ \text{kg} \times 5 \times 10^{6}\ \text{m s}^{-1}} = 6.4 \times 10^{-15}\ \text{m}. \tag{1.64}$$

This distance is comparable to the size of a nucleus.

(b) The position uncertainty for the person is

$$\Delta x \geq \frac{\hbar}{2\Delta p} \simeq \frac{\hbar}{2mv} = \frac{1.05 \times 10^{-34}\ \text{J s}}{2 \times 50\ \text{kg} \times 2\ \text{m s}^{-1}} = 0.5 \times 10^{-36}\ \text{m}. \tag{1.65}$$

An uncertainty of this magnitude is beyond human detection; therefore, it can be neglected. The accuracy of the person's position is limited only by the uncertainties induced by the device used in the measurement. So the position and momentum uncertainties are important for microscopic systems, but negligible for macroscopic systems.

1.5.2 Probabilistic Interpretation

In quantum mechanics the state (or one of the states) of a particle is described by a *wave function* $\psi(\vec{r}, t)$ corresponding to the de Broglie wave of this particle; so $\psi(\vec{r}, t)$ describes the wave properties of a particle. As a result, when discussing quantum effects, it is suitable to use the amplitude function, ψ, whose square modulus, $|\psi|^2$, is equal to the intensity of the wave associated with this quantum effect. The intensity of a wave at a given point in space is proportional to the probability of finding, at that point, the material particle that corresponds to the wave.

In 1927 Born interpreted $|\psi|^2$ as the *probability density* and $|\psi(\vec{r}, t)|^2 d^3r$ as the probability, $dP(\vec{r}, t)$, of finding a particle at time t in the volume element d^3r located between $\vec{r}$ and $\vec{r} + d\vec{r}$:

$$|\psi(\vec{r}, t)|^2 d^3r = dP(\vec{r}, t), \tag{1.66}$$

where $|\psi|^2$ has the dimensions of $[\text{Length}]^{-3}$. If we integrate over the entire space, we are certain that the particle is somewhere in it. Thus, the total probability of finding the particle somewhere in space must be equal to one:

$$\int_{\text{all space}} |\psi(\vec{r}, t)|^2 d^3r = 1. \tag{1.67}$$

The main question now is, how does one determine the wave function ψ of a particle? The answer to this question is given by the theory of quantum mechanics, where ψ is determined by the Schrödinger equation (Chapters 3 and 4).

1.6 Bohr's Atomic Model, Atomic Transitions and Spectroscopy

Besides failing to explain blackbody radiation, the Compton, photoelectric, and pair production effects and the wave–particle duality, classical physics also fails to account for many other

phenomena at the microscopic scale. In this section we consider another area where classical physics breaks down—the atom. Experimental observations reveal that atoms exist as stable, bound systems that have *discrete* numbers of energy levels. Classical physics, however, states that any such bound system must have a continuum of energy levels.

1.6.1 Rutherford Planetary Model of the Atom

After his experimental discovery of the atomic nucleus in 1911, Rutherford proposed a model in an attempt to explain the properties of the atom. Inspired by the orbiting motion of the planets around the sun, Rutherford considered the atom to consist of electrons orbiting around a positively charged massive center, the nucleus. It was soon recognized that, within the context of *classical physics*, this model suffers from two serious deficiencies: (a) atoms are *unstable* and (b) atoms radiate energy over a *continuous* range of frequencies.

The first deficiency results from the application of Maxwell's electromagnetic theory to Rutherford's model: as the electron orbits around the nucleus, it accelerates and hence radiates energy. It must therefore lose energy during its orbiting motion. The radius of the orbit should then decrease continuously (spiral motion) until the electron collapses onto the nucleus; the typical time for such a collapse is about 10^{-8} s. Second, since the frequency of the radiated energy is the same as the orbiting frequency, and as the electron orbit collapses, its orbiting frequency increases *continuously*. Thus, the spectrum of the radiation emitted by the atom should be continuous. These two conclusions completely disagree with experiment, since atoms are *stable* and radiate energy over *discrete* frequency ranges.

1.6.2 Bohr Model of the Hydrogen Atom

Combining Rutherford's planetary model, Planck's quantum hypothesis, and Einstein's photon concept, Bohr proposed in 1913 a model that gives an accurate account of the observed spectrum of the hydrogen atom as well as a convincing explanation for its stability.

Bohr assumed, as in Rutherford's model, that each atom's electron moves in an orbit around the nucleus under the influence of the electrostatic attraction of the nucleus; circular or elliptic orbits are allowed by classical mechanics. For simplicity, Bohr considered only circular orbits, and introduced three, quite arbitrary, assumptions (postulates) which violate classical physics but which turned out to be immensely successful in explaining many properties of the hydrogen atom:

1. Instead of a *continuum* of orbits, which are possible in classical mechanics, only a *discrete* set of circular *stable* orbits, called *stationary states*, are allowed. Atoms can exist only in certain stable states with definite energies: E_1, E_2, E_3, etc.

2. The allowed (stationary) orbits correspond to those for which the orbital angular momentum of the electron is an *integer multiple* of $\hbar$ ($\hbar = h/2\pi$):

$$L = n\hbar. \tag{1.68}$$

   This relation is known as the *Bohr quantization rule* of the angular momentum.

3. As long as an electron remains in a stationary orbit, it does not radiate electromagnetic energy. Emission or absorption of radiation can take place only when an electron

jumps from one allowed orbit to another. The radiation corresponding to the electron's transition from an orbit of energy E_n to another E_m is carried out by a photon of energy

$$h\nu = E_n - E_m. \tag{1.69}$$

So an atom may emit (or absorb) radiation by having the electron jump to a lower (or higher) orbit.

From the very outset, Bohr realized that, since there was no way to explain the stability of atoms within classical physics, he had to go beyond its scope; this had encouraged him to posit three postulates that violated classical physics and proposed his *quantized* shell model for the hydrogen atom. In what follows we are going to apply Bohr's postulates to the hydrogen atom. We want to provide a quantitative description of its energy levels and its spectroscopy.

1.6.2.1 Energy Levels of the Hydrogen Atom

Let us see how Bohr's quantization condition (1.68) leads to a discrete set of energies E_n and radii r_n. When the electron of the hydrogen atom moves in a *circular* orbit, the application of Newton's second law to the electron yields $F = m_e a_r = m_e v^2/r$. Since the only force[55] acting on the electron is the electrostatic force applied on it by the proton, we can equate the electrostatic force to the centripetal force and obtain

$$\frac{e^2}{4\pi\varepsilon_0 r^2} = m_e \frac{v^2}{r}. \tag{1.70}$$

Now, assumption (1.68) yields

$$L = m_e v r = n\hbar, \tag{1.71}$$

hence $m_e v^2/r = n^2\hbar^2/(m_e r^3)$, which when combined with (1.70) yields $e^2/(4\pi\varepsilon_0 r^2) = n^2\hbar^2/(m_e r^3)$; this relation in turn leads to a quantized expression for the radius:

$$\boxed{r_n = \left(\frac{4\pi\varepsilon_0\hbar^2}{m_e e^2}\right) n^2 = n^2 a_0,} \tag{1.72}$$

where

$$a_0 = \frac{4\pi\varepsilon_0\hbar^2}{m_e e^2} \tag{1.73}$$

is the *Bohr radius*, $a_0 = 0.053$ nm. The speed of the orbiting electron can be obtained from (1.71) and (1.72):

$$\boxed{v_n = \frac{n\hbar}{m_e r_n} = \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{n\hbar}.} \tag{1.74}$$

Note that the ratio between the speed of the electron in the first Bohr orbit, v_1, and the speed of light is equal to a dimensionless constant α, known as the *fine structure constant*:

$$\alpha = \frac{v_1}{c} = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{\hbar c} = \frac{1}{137} \Rightarrow v_1 = \alpha c = \frac{3\times 10^8\ \mathrm{m\,s^{-1}}}{137} \simeq 2.19\times 10^6\ \mathrm{m\,s^{-1}}. \tag{1.75}$$

[55] At the atomic scale, gravity has no measurable effect. The gravitational force between the hydrogen's proton and electron, $F_G = (Gm_e m_p)/r^2$, is negligible compared to the electrostatic force $F_e = e^2/(4\pi\varepsilon_0 r^2)$, since $F_G/F_e = Gm_e m_p(4\pi\varepsilon_0/e^2) \simeq 10^{-40}$.

As for the total energy of the electron, it is given by

$$E = \frac{1}{2}m_e v^2 - \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r}; \tag{1.76}$$

in deriving this relation, we have assumed that the nucleus, i.e., the proton, is infinitely heavy compared with the electron and hence it can be considered at rest; that is, the energy of the electron–proton system consists of the kinetic energy of the electron plus the electrostatic potential energy. From (1.70) we see that the kinetic energy, $\frac{1}{2}m_e v^2$, is equal to $\frac{1}{2}e^2/(4\pi\varepsilon_0 r)$, which when inserted into (1.76) leads to

$$E = -\frac{1}{2}\left(\frac{e^2}{4\pi\varepsilon_0 r}\right). \tag{1.77}$$

This equation shows that the electron circulates in an orbit of radius r with a kinetic energy equal to minus one half the potential energy (this result is the well known Virial theorem of classical mechanics). Substituting r_n of (1.72) into (1.77), we obtain

$$\boxed{E_n = -\frac{e^2}{8\pi\varepsilon_0}\frac{1}{r_n} = -\frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\frac{1}{n^2} = -\frac{\mathcal{R}}{n^2},} \tag{1.78}$$

known as the *Bohr energy*, where $\mathcal{R}$ is the *Rydberg constant*:

$$\mathcal{R} = \frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 = 13.6\ \text{eV}. \tag{1.79}$$

The energy E_n of each state of the atom is determined by the value of the quantum number n. The *negative* sign of the energy (1.78) is due to the *bound* state nature of the atom. That is, states with negative energy $E_n < 0$ correspond to bound states.

The structure of the atom's energy spectrum as given by (1.78) is displayed in Figure 1.13 (where, by convention, the energy levels are shown as horizontal lines). As n increases, the energy level separation decreases rapidly. Since n can take all integral values from $n = 1$ to $n = +\infty$, the energy spectrum of the atom contains an infinite number of discrete energy levels. In the ground state ($n = 1$), the atom has an energy $E_1 = -\mathcal{R}$ and a radius a_0. The states $n = 2, 3, 4, \ldots$ correspond to the *excited states* of the atom, since their energies are greater than the ground state energy.

When the quantum number n is very large, $n \to +\infty$, the atom's radius r_n will also be very large but the energy values go to zero, $E_n \to 0$. This means that the proton and the electron are infinitely far away from one another and hence they are no longer bound; the atom is ionized. In this case there is no restriction on the amount of kinetic energy the electron can take, for it is free. This situation is represented in Figure 1.13 by the *continuum* of positive energy states, $E_n > 0$.

Recall that in deriving (1.72) and (1.78) we have neglected the mass of the proton. If we include it, the expressions (1.72) and (1.78) become

$$\boxed{r_n = \frac{4\pi\varepsilon_0\hbar^2}{\mu e^2}n^2 = \left(1 + \frac{m_e}{m_p}\right)a_0 n^2, \quad E_n = -\frac{\mu}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\frac{1}{n^2} = -\frac{1}{1 + m_e/m_p}\frac{\mathcal{R}}{n^2},} \tag{1.80}$$

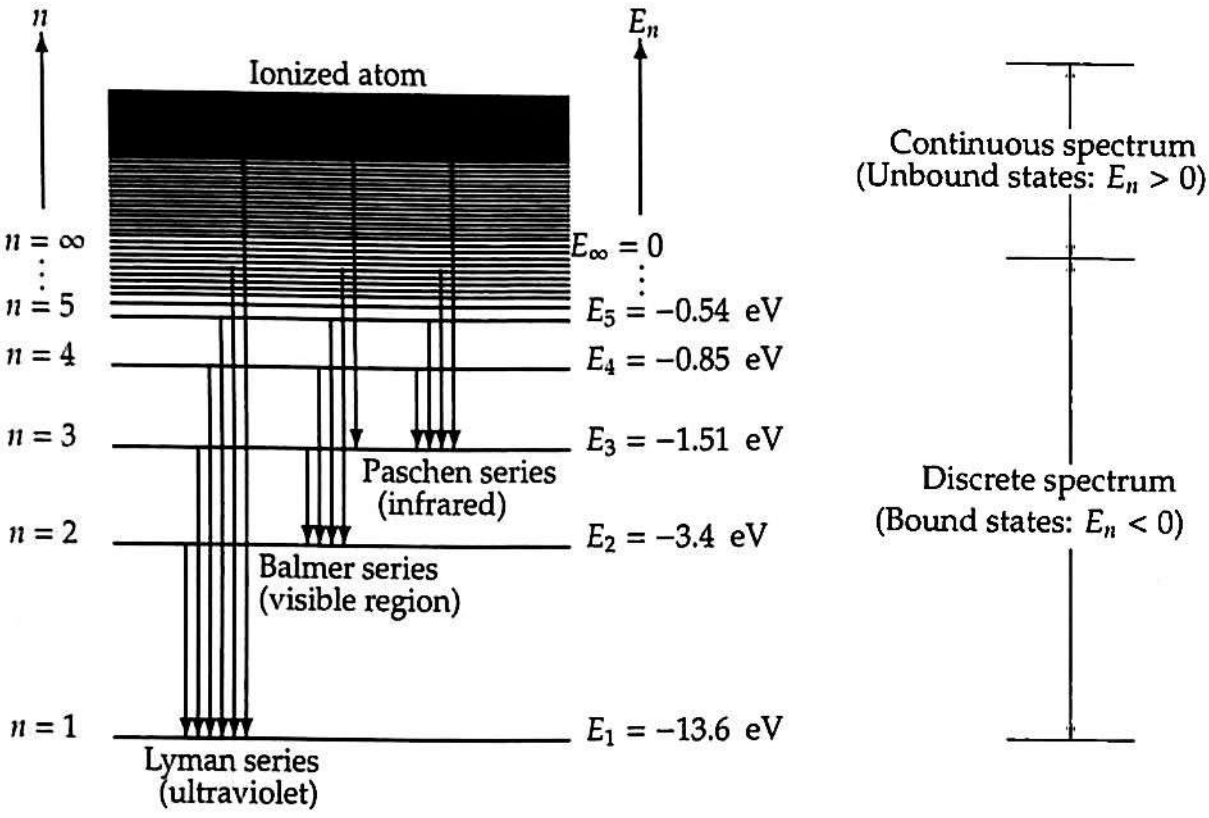


Figure 1.13: Energy levels and transitions between them for the hydrogen atom.

where $\mu = m_p m_e/(m_p+m_e) = m_e/(1+m_e/m_p)$ is the reduced mass of the proton–electron system.

We should note that r_n and E_n of (1.80), which were derived for the hydrogen atom, can be generalized to hydrogen-like ions where all electrons save one are removed. To obtain the radius and energy of a single electron orbiting a fixed nucleus of Z protons, we need simply to replace e^2 in (1.80) by Ze^2,

$$\boxed{r_n = \left(1+\frac{m_e}{M}\right)\frac{a_0}{Z}n^2, \qquad E_n = -\frac{Z^2}{1+m_e/M}\frac{\mathcal{R}}{n^2},} \tag{1.81}$$

where M is the mass of the nucleus; when $m_e/M \ll 1$ we can just drop the term m_e/M.

Remarks

We should note that, in spite of its enormous successes, the Bohr model suffers from a number of flaws. We may mention, for instance, the fact that the Bohr model calculations of the energy and radii work only on single-electron atoms or ions; the model provides no method for calculating the energy levels or radii of multi-electron atoms or ions. Another limitation of the Bohr model is that it is essentially one-dimensional: it uses only one quantum number (the principle quantum number "n") to describe the distribution of electrons in an atom. This is fact is evident from Eqs. (1.72) and (1.78) which show that the radius of the electron's orbit and the energy depend on n only, $r_n = a_0 n^2$ and $E_n = -\mathcal{R}/n^2$. In contrast, we will see in Chapter 6 that in the Schrödinger formulation of quantum mechanics, the state of an electron in an atom is described by means of a wave function that depends on three quantum numbers (n, l, m) called the principal, orbital, and magnetic quantum numbers,

respectively; this is expected since three coordinates (or three quantum numbers) are required to describe an electron within a three-dimensional space. Another major problem with the Bohr model is that it treats electrons as point-like particles that rotate around the nucleus on fixed *"sharp"* circular orbits, like planets orbiting around the sun (see Figure 6.5-(Left)); this is a gross oversimplification of reality as one cannot find the position of the electron with infinite accuracy. As we will see in Chapter 6, the Schrödinger or quantum mechanical model treats electrons as three-dimensional standing waves that spread around the nucleus like *clouds* (see Figure 6.5-(Right)). In sharp contrast to the Bohr model, only a probabilistic outcome is possible when it comes to measuring the position of an atom's electron within the framework of quantum mechanics; the probability of finding the electron is highest in those regions where the electronic cloud is densest (i.e., at the crests or antinodes of the standing wave) and smallest where the cloud is thinnest (i.e., at the nodes).

Connection between de Broglie's Hypothesis and Bohr's Quantization Condition

The Bohr quantization condition (1.68), $L = n\hbar$, can be viewed as a manifestation of de Broglie's hypothesis (1.50), $\lambda = h/(mv)$. Since an atom's electron is represented by a circular standing wave and for the wave associated with the electron to be a standing wave, the circumference of the electron's orbit must be equal to an *integral* multiple of the electron's wavelength:

$$2\pi r = n\lambda \qquad (n = 1, 2, 3, \ldots). \tag{1.82}$$

This relation can be reduced to (1.68) or to (1.71), provided that we make use of de Broglie's relation, $\lambda = h/p = h/(m_e v)$. That is, inserting $\lambda = h/(m_e v)$ into (1.82) and using the fact that the electron's orbital angular momentum is $L = m_e vr$, we have

$$2\pi r = n\lambda = n\frac{h}{m_e v} \quad \Longrightarrow \quad m_e vr = n\frac{h}{2\pi} \quad \Longrightarrow \quad L = n\hbar, \tag{1.83}$$

which is identical with Bohr's quantization condition (1.68). In essence, this condition states that the only allowed orbits for the electron are those whose circumferences are equal to integral multiples of the de Broglie wavelength. For example, in the case of a hydrogen atom and as illustrated in Figure (1.14), the circumference of the electron's orbit is equal to λ_1 when the electron is in the ground state ($n = 1$); it is equal to $2\lambda_2$ when the electron is in the first excited state ($n = 2$); equal to $3\lambda_3$ when it is in its second excited state ($n = 3$); and so on. That is, since $r_n = n^2 a_0$ where $a_0 = 0.053$ nm, the circumference of each orbit is equal to

$$2\pi r_n \equiv 2\pi n^2 a_0 = n\lambda_n \quad \Longrightarrow \quad \lambda_n = 2\pi n a_0 \qquad (n = 1, 2, 3, \ldots). \tag{1.84}$$

For instance, for $n = 1$, we have $\lambda_1 = 2\pi a_0 = 6.28a_0 = 0.333$ nm; for $n = 2$, $\lambda_2 = 4\pi a_0 = 12.57a_0 = 0.666$ nm; for $n = 3$, $\lambda_3 = 6\pi a_0 = 18.85a_0 = 0.999$ nm; for $n = 4$, $\lambda_4 = 8\pi a_0 = 25.13a_0 = 1.332$ nm; for $n = 5$, $\lambda_5 = 10\pi a_0 = 31.42a_0 = 1.665$ nm; and so on.

S Since the wavelength of an object is inversely proportional to its mass (as per de Broglie's relation $\lambda = h/(mv)$), small objects have wavelengths that are of the same size as them and massive objects have negligible wavelengths. Thus, while the wavelike aspect of small objects is important, the wavelike behavior of massive objects is not.

1.6.2.2 Spectroscopy of the Hydrogen Atom (Rydberg Formula)

Having specified the energy spectrum of the hydrogen atom, let us now study its spectroscopy. In sharp contrast to the continuous nature of the spectral distribution of the radiation emitted

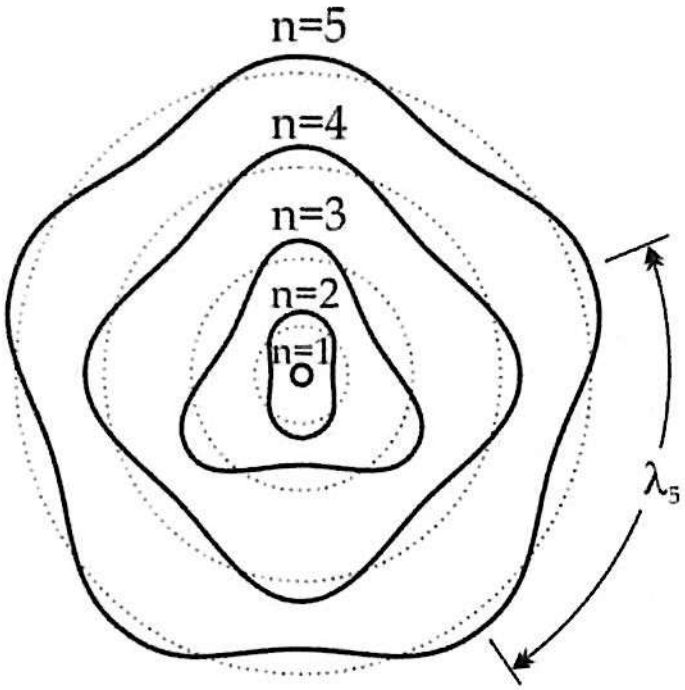


Figure 1.14: Illustration of the five lowest orbits of the electron in a hydrogen atom where the electron is represented by a wave as per de Broglie's matter-wave hypothesis.

by glowing solid objects, the radiation emitted or absorbed by a gas displays a *discrete* spectrum distribution. When subjecting a gas to an electric discharge (or to a flame), the radiation emitted from the excited atoms of the gas discharge consists of a few sharp lines (bright lines of pure color, with darkness in between). A major success of Bohr's model lies in its ability to predict accurately the *sharpness* of the spectral lines emitted or absorbed by the atom. The model shows clearly that these discrete lines correspond to the sharply defined energy levels of the atom. The radiation emitted from the atom results from the transition of the electron from an allowed state n to another m; this radiation has a well defined (sharp) frequency ν which can be obtained from the Bohr energy expression (1.78)

$$\boxed{h\nu = E_n - E_m = \mathcal{R}\left(\frac{1}{m^2} - \frac{1}{n^2}\right).} \tag{1.85}$$

So when an electron absorbs a photon (of energy $h\nu$), its wavelength increases and in the process acquiring more energy to the point of jumping to a higher orbit due to excitation; conversely, when an electron looses energy (by giving away a photon), its wavelength decreases and it will migrate to a lower orbit. For instance, the *Lyman series*, which corresponds to the emission of *ultraviolet* radiation, is due to transitions from excited states $n = 2, 3, 4, 5, \ldots$ to the ground state $n = 1$ (see Figure 1.13):

$$h\nu_L = E_n - E_1 = \mathcal{R}\left(\frac{1}{1^2} - \frac{1}{n^2}\right) \qquad (n > 1). \tag{1.86}$$

Another transition series, the *Balmer series*, is due to transitions to the first excited state ($n = 2$):

$$h\nu_B = E_n - E_2 = \mathcal{R}\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \qquad (n > 2). \tag{1.87}$$

The atom emits *visible* radiation as a result of the Balmer transitions. As shown in Table 1.1 and Figure 1.13, other series are Paschen, $n \to 3$ with $n > 3$; Brackett, $n \to 4$ with $n > 4$; Pfund, $n \to 5$ with $n > 5$; , Humprefys, $n \to 6$ with $n > 5$; and so on. They correspond to the emission of *infrared* radiation. We should mention that the results obtained from (1.85) are in spectacular agreement with those of experimental spectroscopy.

Remark: The Bohr model gives back the Balmer and Rydberg formulas
At this level, it worth mentioning that equation (1.85), which was derived from the Bohr energy expression (1.78), is quite similar to the formulas (1.1) and (1.2) that Balmer and Rydberg had obtained back in 1885 and 1888, respectively, using a heuristic method which was based on a numerical, trial-and-error approach. Since they were not able to obtain their formulas from classical physics (due to its invalidity at the microscopic level), they worked their way backwards by starting from the experimental data and looked for a mathematical formula that can reproduce the data.

Interaction of radiation with matter
So far in this chapter, we have seen that when a photon (radiation) passes through matter, it will interact with it according to one of the following three scenarios:

1. If it comes in contact with an electron that is at rest, it will scatter from it like a corpuscular particle: it will impart a momentum to the electron, it will scatter and continue its travel with the speed of light but with a lower frequency (or higher wavelength). This is the Compton effect.

2. If it comes into contact with an atom's electron, it will interact with the electron according to one of the following two scenarios:

   - If it has enough energy, it will knock the electron completely out of the atom and then vanish, for it transmits all its energy to the electron. This is the photoelectric effect.
   - If its energy $h\nu$ is not sufficient to knock out the electron altogether, it will kick the electron to a higher orbit, provided $h\nu$ is equal to the energy difference between the initial and final orbits: $h\nu = E_n - E_m$. In the process it will transmit all its energy to the electron and then vanish. The atom will be left in an excited state. However, if $h\nu \neq E_n - E_m$, nothing will happen (the photon simply scatters away).

3. If it comes in contact with an atomic nucleus and if its energy is sufficiently high ($h\nu \geq 2m_e c^2$), it will vanish by creating matter: an electron–positron pair will be produced. This is pair production.

Example 1.7 (Positronium's radius and energy spectrum)
Positronium is the bound state of an electron and a positron; it is a short-lived, hydrogen-like atom where the proton is replaced by a positron.

(a) Calculate the energy and radius expressions, E_n and r_n.

(b) Estimate the values of the energies and radii of the three lowest states.

(c) Calculate the frequency and wavelength of the electromagnetic radiation that will just ionize the positronium atom when it is in its first excited state.

Solution

(a) The radius and energy expressions of the positronium can be obtained at once from (1.80) by simply replacing the reduced mass μ with that of the electron–positron system $\mu = m_e m_e/(m_e + m_e) = \frac{1}{2}m_e$:

$$r_n = \left(\frac{8\pi\varepsilon_0\hbar^2}{m_e e^2}\right)n^2, \qquad E_n = -\frac{m_e}{4\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\frac{1}{n^2}. \tag{1.88}$$

We can rewrite r_n and E_n in terms of the Bohr radius, $a_0 = 4\pi\varepsilon_0\hbar^2/(m_e e^2) = 0.053$ nm, and the Rydberg constant, $\mathcal{R} = \frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 = 13.6$ eV, as follows:

$$r_n = 2a_0 n^2, \qquad E_n = -\frac{\mathcal{R}}{2n^2}. \tag{1.89}$$

These are related to the expressions for the hydrogen by $r_{n_{pos}} = 2r_{n_H}$ and $E_{n_{pos}} = \frac{1}{2}E_{n_H}$.

(b) The radii of the three lowest states of the positronium are given by $r_1 = 2a_0 = 0.106$ nm, $r_2 = 8a_0 = 0.424$ nm, and $r_3 = 18a_0 = 0.954$ nm. The corresponding energies are $E_1 = -\frac{1}{2}\mathcal{R} = -6.8$ eV, $E_2 = -\frac{1}{8}\mathcal{R} = -1.7$ eV, and $E_3 = -\frac{1}{18}\mathcal{R} = -0.756$ eV.

(c) Since the energy of the first excited state of the positronium is $E_2 = -1.7$ eV $= -1.7 \times 1.6 \times 10^{-19}$ J $= -2.72 \times 10^{-19}$ J, the energy of the electromagnetic radiation that will just ionize the positronium is equal to $h\nu = E_\infty - E_2 = 0 - (-2.72 \times 10^{-19}\ \text{J}) = 2.72 \times 10^{-19}\ \text{J} = E_{ion}$; hence the frequency and wavelength of the ionizing radiation are given by

$$\nu = \frac{E_{ion}}{h} = \frac{2.72 \times 10^{-19}\ \text{J}}{6.6 \times 10^{-34}\ \text{J s}} = 4.12 \times 10^{14}\ \text{Hz}, \tag{1.90}$$

$$\lambda = \frac{c}{\nu} = \frac{3 \times 10^8\ \text{m s}^{-1}}{4.12 \times 10^{14}\ \text{Hz}} = 7.28 \times 10^{-7}\ \text{m}. \tag{1.91}$$

1.7 Quantization Rules

The ideas that led to successful explanations of blackbody radiation, the photoelectric effect, and the hydrogen's energy levels rest on two quantization rules: (a) the relation (1.12) that Planck postulated to explain the quantization of energy, $E = nh\nu$, and (b) the condition (1.68) that Bohr postulated to account for the quantization of the electron's orbital angular momentum, $L = n\hbar$. A number of attempts were undertaken to understand or interpret these rules. In 1916 Wilson and Sommerfeld offered a scheme that included both quantization rules as special cases. In essence, their scheme, which applies only to systems with coordinates that are periodic in time, consists in *quantizing the action variable*, $J = \oint p\,dq$, of classical mechanics:

$$\oint p\,dq = nh \qquad (n = 0, 1, 2, 3, \ldots), \tag{1.92}$$

where n is a quantum number, p is the momentum conjugate associated with the coordinate q; the closed integral $\oint$ is taken over one period of q. This relation is known as the Wilson–Sommerfeld quantization rule.

Wilson–Sommerfeld quantization rule and Planck's quantization relation

In what follows we are going to show how the Wilson–Sommerfeld rule (1.92) leads to Planck's quantization relation $E = nh\nu$. For an illustration, consider a one-dimensional harmonic oscillator where a particle of mass m oscillates harmonically between $-a \leq x \leq a$; its classical energy is given by

$$E(x,p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2; \tag{1.93}$$

hence $p(E,x) = \pm\sqrt{2mE - m^2\omega^2x^2}$. At the turning points, $x_{min} = -a$ and $x_{max} = a$, the energy is purely potential: $E = V(\pm a) = \frac{1}{2}m\omega^2a^2$; hence $a = \sqrt{2E/(m\omega^2)}$. Using $p(E,x) = \pm\sqrt{2mE - m^2\omega^2x^2}$ and from symmetry considerations, we can write the action as

$$\oint p\,dx = 2\int_{-a}^{a}\sqrt{2mE - m^2\omega^2x^2}dx = 4m\omega\int_0^a\sqrt{a^2 - x^2}dx. \tag{1.94}$$

The change of variables $x = a\sin\theta$ leads to

$$\int_0^a\sqrt{a^2 - x^2}dx = a^2\int_0^{\pi/2}\cos^2\theta\,d\theta = \frac{a^2}{2}\int_0^{\pi/2}(1 + \cos 2\theta)d\theta = \frac{\pi a^2}{4} = \frac{\pi E}{2m\omega^2}. \tag{1.95}$$

Since $\omega = 2\pi\nu$, where ν is the frequency of oscillations, we have

$$\oint p\,dx = \frac{2\pi E}{\omega} = \frac{E}{\nu}. \tag{1.96}$$

Inserting (1.96) into (1.92), we end up with the Planck quantization rule $E = nh\nu$, i.e.,

$$\oint p\,dx = nh \quad\Longrightarrow\quad \frac{E}{\nu} = nh \quad\Longrightarrow\quad E_n = nh\nu. \tag{1.97}$$

We can interpret this relation as follows. From classical mechanics, we know that the motion of a mass subject to harmonic oscillations is represented in the xp phase space by a continuum of ellipses whose areas are given by $\oint p\,dx = E/\nu$, because the integral $\oint p(x)\,dx$ gives the area enclosed by the closed trajectory of the particle in the xp phase space. The condition (1.92) or (1.97) provides a mechanism for selecting, from the continuum of the oscillator's energy values, only those energies E_n for which the areas of the contours $p(x,E_n) = \sqrt{2m\,(E_n - V(x))}$ are equal to nh with $n = 0, 1, 2, 3, \ldots$. That is, the only allowed states of oscillation are those represented in the phase space by a series of ellipses with "quantized" areas $\oint p\,dx = nh$. Note that the area between two successive states is equal to h: $\oint p(x,E_{n+1})\,dx - \oint p(x,E_n)\,dx = h$.

This simple calculation shows that the *Planck rule for energy quantization is equivalent to the quantization of action.*

Wilson–Sommerfeld quantization rule and Bohr's quantization condition

Let us now show how the Wilson–Sommerfeld rule (1.92) leads to Bohr's quantization condition (1.68). For an electron moving in a circular orbit of radius r, it is suitable to use polar coordinates (r,φ). The action $J = \oint p\,dq$, which is expressed in Cartesian coordinates by the linear momentum p and its conjugate variable x, is characterized in polar coordinates by the

orbital angular momentum L and its conjugate variable φ, the polar angle, where φ is periodic in time. That is, $J = \oint p\,dq$ is given in polar coordinates by $\int_0^{2\pi} L\,d\varphi$. In this case (1.92) becomes

$$\int_0^{2\pi} L\,d\varphi = nh. \tag{1.98}$$

For spherically symmetric potentials—as it is the case here where the electron experiences the proton's Coulomb potential—the angular momentum L is a constant of the motion. Hence (1.98) shows that angular momentum can change only in integral units of $\hbar$:

$$L\int_0^{2\pi} d\varphi = nh \qquad \Longrightarrow \qquad L = n\frac{h}{2\pi} = n\hbar, \tag{1.99}$$

which is identical with the Bohr quantization condition (1.68). This calculation also shows that the *Bohr quantization is equivalent to the quantization of action.* As stated above (1.83), the Bohr quantization condition (1.68) has the following physical meaning: while orbiting the nucleus, the electron moves only in well specified orbits, orbits with circumferences equal to integral multiples of the de Broglie wavelength.

Note that the Wilson–Sommerfeld quantization rule (1.92) does not tell us how to calculate the energy levels of non-periodic systems; it applies only to systems which are periodic. On a historical note, the quantization rules of Planck and Bohr have dominated quantum physics from 1900 to 1925; the quantum physics of this period is known as the "old quantum theory." The success of these quantization rules, as measured by the striking agreement of their results with experiment, gave irrefutable evidence for the quantization hypothesis of all material systems and constituted a triumph of the "old quantum theory." In spite of their quantitative success, these quantization conditions suffer from a serious inconsistency: they do not originate from a theory, they were postulated rather arbitrarily.

1.8 Wave Packets

At issue here is how to describe a particle within the context of quantum mechanics. As quantum particles jointly display particle and wave features, we need to look for a mathematical scheme that can embody them simultaneously.

In classical physics, a particle is *well localized* in space, for its position and velocity can be calculated simultaneously to arbitrary precision. As for quantum mechanics, it describes a material particle by a *wave function* corresponding to the matter wave associated with the particle (de Broglie's conjecture). Wave functions, however, depend on the *whole* space; hence they *cannot be localized.* If the wave function is made to vanish everywhere except in the neighborhood of the particle or the neighborhood of the "classical trajectory," it can then be used to describe the dynamics of the particle. That is, a particle which is localized within a certain region of space can be described by a matter wave whose amplitude is large in that region and zero outside it. This matter wave must then be *localized* around the region of space within which the particle is confined.

A *localized* wave function is called a *wave packet.* A wave packet therefore consists of a group of waves of slightly different wavelengths, with phases and amplitudes so chosen that they interfere constructively over a small region of space and destructively elsewhere. Not

only are wave packets useful in the description of "isolated" particles that are confined to a certain spatial region, they also play a key role in understanding the connection between quantum mechanics and classical mechanics. The wave packet concept therefore represents a unifying mathematical tool that can cope with and embody nature's particle-like behavior and also its wave-like behavior.

1.8.1 Localized Wave Packets

Localized wave packets can be constructed by superposing, in the same region of space, waves of slightly different wavelengths, but with phases and amplitudes chosen to make the superposition constructive in the desired region and destructive outside it. Mathematically, we can carry out this superposition by means of *Fourier transforms*. For simplicity, we are going to consider a one-dimensional wave packet; this packet is intended to describe a "classical" particle confined to a one-dimensional region, for instance, a particle moving along the x-axis. We can construct the packet $\psi(x,t)$ by superposing plane waves (propagating along the x-axis) of different frequencies (or wavelengths):

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{i(kx-\omega t)} dk; \tag{1.100}$$

$\phi(k)$ is the amplitude of the wave packet.

In what follows we want to look at the form of the packet at a given time; we will deal with the time evolution of wave packets later. Choosing this time to be $t = 0$ and abbreviating $\psi(x,0)$ by $\psi_0(x)$, we can reduce (1.100) to

$$\psi_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{ikx} dk, \tag{1.101}$$

where $\phi(k)$ is the Fourier transform of $\psi_0(x)$,

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \psi_0(x) e^{-ikx} dx. \tag{1.102}$$

The relations (1.101) and (1.102) show that $\phi(k)$ determines $\psi_0(x)$ and vice versa. The packet (1.101), whose form is determined by the x-dependence of $\psi_0(x)$, does indeed have the required property of localization: $|\psi_0(x)|$ peaks at $x = 0$ and vanishes far away from $x = 0$. On the one hand, as $x \to 0$ we have $e^{ikx} \to 1$; hence the waves of different frequencies interfere constructively (i.e., the various k-integrations in (1.101) add constructively). On the other hand, far away from $x = 0$ (i.e., $|x| \gg 0$) the phase e^{ikx} goes through many periods leading to violent oscillations, thereby yielding destructive interference (i.e., the various k-integrations in (1.101) add up to zero). This implies, in the language of Born's probabilistic interpretation, that the particle has a greater probability of being found near $x = 0$ and a scant chance of being found far away from $x = 0$. The same comments apply to the amplitude $\phi(k)$ as well: $\phi(k)$ peaks at $k = 0$ and vanishes far away. Figure 1.15 displays a typical wave packet that has the required localization properties we have just discussed.

In summary, the particle is represented not by a single de Broglie wave of well-defined frequency and wavelength, but by a wave packet that is obtained by adding a large number of waves of different frequencies.

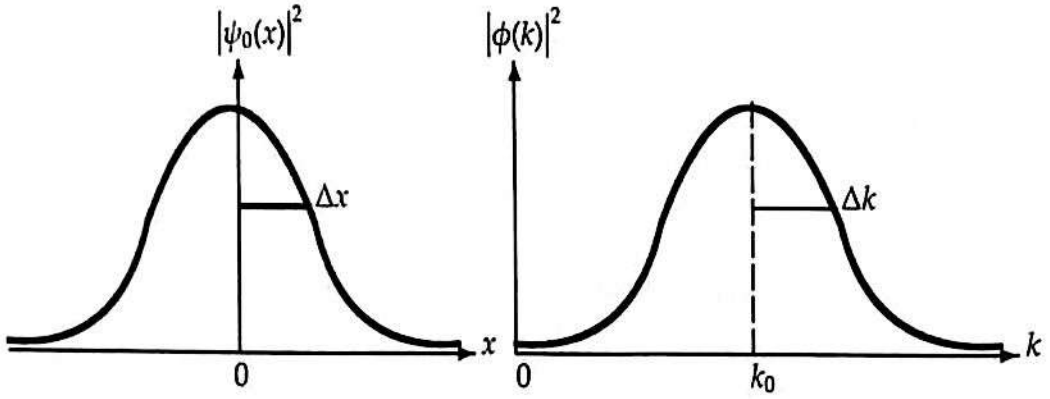


Figure 1.15: Two localized wave packets: $\psi_0(x) = (2/\pi a^2)^{1/4} e^{-x^2/a^2} e^{ik_0 x}$ and $\phi(k) = (a^2/2\pi)^{1/4} e^{-a^2(k-k_0)^2/4}$; they peak at $x = 0$ and $k = k_0$, respectively, and vanish far away.

The physical interpretation of the wave packet is obvious: $\psi_0(x)$ is the wave function or probability amplitude for finding the particle at position x; hence $|\psi_0(x)|^2$ gives the probability density for finding the particle at x, and $P(x)\,dx = |\psi_0(x)|^2 dx$ gives the probability of finding the particle between x and $x+dx$. What about the physical interpretation of $\phi(k)$? From (1.101) and (1.102) it follows that

$$\int_{-\infty}^{+\infty} |\psi_0(x)|^2 dx = \int_{-\infty}^{+\infty} |\phi(k)|^2 dk; \tag{1.103}$$

then if $\psi(x)$ is normalized so is $\phi(k)$, and vice versa. Thus, the function $\phi(k)$ can be interpreted most naturally, like $\psi_0(x)$, as a probability amplitude for measuring a wave vector k for a particle in the state $\phi(k)$. Moreover, while $|\phi(k)|^2$ represents the probability density for measuring k as the particle's wave vector, the quantity $P(k)\,dk = |\phi(k)|^2 dk$ gives the probability of finding the particle's wave vector between k and $k + dk$.

We can extract information about the particle's motion by simply expressing its corresponding matter wave in terms of the particle's energy, E, and momentum, p. Using $k = p/\hbar$, $dk = dp/\hbar$, $E = \hbar\omega$ and redefining $\tilde{\phi}(p) = \phi(k)/\sqrt{\hbar}$, we can rewrite (1.100) to (1.102) as follows:

$$\psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \tilde{\phi}(p) e^{i(px-Et)/\hbar} dp, \tag{1.104}$$

$$\psi_0(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \tilde{\phi}(p) e^{ipx/\hbar} dp, \tag{1.105}$$

$$\tilde{\phi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi_0(x) e^{-ipx/\hbar} dx, \tag{1.106}$$

where $E(p)$ is the total energy of the particle described by the wave packet $\psi(x,t)$ and $\tilde{\phi}(p)$ is the momentum amplitude of the packet.

In what follows we are going to illustrate the basic ideas of wave packets on a simple, instructive example: the Gaussian and square wave packets.

Example 1.8 (Gaussian and square wave packets)

(a) Find $\psi(x, 0)$ for a Gaussian wave packet $\phi(k) = A \exp\left[-a^2(k-k_0)^2/4\right]$, where A is a

normalization factor to be found. Calculate the probability of finding the particle in the region $-a/2 \leq x \leq a/2$.

(b) Find $\phi(k)$ for a square wave packet $\psi_0(x) = \begin{cases} Ae^{ik_0x}, & |x| \leq a, \\ 0, & |x| > a. \end{cases}$

Find the factor A so that $\psi(x)$ is normalized.

Solution

(a) The normalization factor A is easy to obtain:

$$1 = \int_{-\infty}^{+\infty} |\phi(k)|^2 dk = |A|^2 \int_{-\infty}^{+\infty} \exp\left[-\frac{a^2}{2}(k-k_0)^2\right] dk, \tag{1.107}$$

which, by using a change of variable $z = k - k_0$ and using the integral $\int_{-\infty}^{+\infty} e^{-a^2z^2/2}dz = \sqrt{2\pi}/a$, leads at once to $A = \sqrt{a/\sqrt{2\pi}} = [a^2/(2\pi)]^{1/4}$. Now, the wave packet corresponding to

$$\phi(k) = \left(\frac{a^2}{2\pi}\right)^{1/4} \exp\left[-\frac{a^2}{4}(k-k_0)^2\right] \tag{1.108}$$

is

$$\psi_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{ikx} dk = \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{2\pi}\right)^{1/4} \int_{-\infty}^{+\infty} e^{-a^2(k-k_0)^2/4 + ikx} dk. \tag{1.109}$$

To carry out the integration, we need simply to rearrange the exponent's argument as follows:

$$-\frac{a^2}{4}(k-k_0)^2 + ikx = -\left[\frac{a}{2}(k-k_0) - \frac{ix}{a}\right]^2 - \frac{x^2}{a^2} + ik_0x. \tag{1.110}$$

The introduction of a new variable $y = a(k-k_0)/2 - ix/a$ yields $dk = 2dy/a$, and when combined with (1.109) and (1.110), this leads to

$$\begin{aligned} \psi_0(x) &= \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{2\pi}\right)^{1/4} \int_{-\infty}^{+\infty} e^{-x^2/a^2} e^{ik_0x} e^{-y^2} \left(\frac{2}{a} dy\right) \\ &= \frac{1}{\sqrt{\pi}} \left(\frac{2}{\pi a^2}\right)^{1/4} e^{-x^2/a^2} e^{ik_0x} \int_{-\infty}^{+\infty} e^{-y^2} dy. \end{aligned} \tag{1.111}$$

Since $\int_{-\infty}^{+\infty} e^{-y^2} dy = \sqrt{\pi}$, this expression becomes

$$\psi_0(x) = \left(\frac{2}{\pi a^2}\right)^{1/4} e^{-x^2/a^2} e^{ik_0x}, \tag{1.112}$$

where e^{ik_0x} is the phase of $\psi_0(x)$; $\psi_0(x)$ is an oscillating wave with wave number k_0 modulated by a Gaussian envelope centered at the origin. We will see later that the phase factor e^{ik_0x} has real physical significance. The wave function $\psi_0(x)$ is complex, as necessitated by quantum mechanics. Note that $\psi_0(x)$, like $\phi(k)$, is normalized. Moreover, equations (1.108) and (1.112) show that the Fourier transform of a Gaussian wave packet is also a Gaussian wave packet.

The probability of finding the particle in the region $-a/2 \le x \le a/2$ can be obtained at once from (1.112):

$$P = \int_{-a/2}^{+a/2} |\psi_0(x)|^2 dx = \sqrt{\frac{2}{\pi a^2}} \int_{-a/2}^{+a/2} e^{-2x^2/a^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-1}^{+1} e^{-z^2/2} dz \simeq \frac{2}{3},$$

where we have used the change of variable $z = 2x/a$.

(b) The normalization of $\psi_0(x)$ is straightforward:

$$1 = \int_{-\infty}^{+\infty} |\psi_0(x)|^2 dx = |A|^2 \int_{-a}^{a} e^{-ik_0 x} e^{ik_0 x} dx = |A|^2 \int_{-a}^{a} dx = 2a|A|^2; \tag{1.113}$$

hence $A = 1/\sqrt{2a}$. The Fourier transform of $\psi_0(x)$ is

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \psi_0(x) e^{-ikx} dx = \frac{1}{2\sqrt{\pi a}} \int_{-a}^{a} e^{ik_0 x} e^{-ikx} dx = \frac{1}{\sqrt{\pi a}} \frac{\sin[(k-k_0)a]}{k-k_0}. \tag{1.114}$$

1.8.2 Wave Packets and the Uncertainty Relations

We want to show here that the width of a wave packet $\psi_0(x)$ and the width of its amplitude $\phi(k)$ are not independent; they are correlated by a reciprocal relationship. As it turns out, the reciprocal relationship between the widths in the x and k spaces has a direct connection to Heisenberg's uncertainty relation.

For simplicity, let us illustrate the main ideas on the Gaussian wave packet treated in the previous example (see (1.108) and (1.112)):

$$\psi_0(x) = \left(\frac{2}{\pi a^2}\right)^{1/4} e^{-x^2/a^2} e^{ik_0 x}, \qquad \phi(k) = \left(\frac{a^2}{2\pi}\right)^{1/4} e^{-a^2(k-k_0)^2/4}. \tag{1.115}$$

As displayed in Figure 1.15, $|\psi_0(x)|^2$ and $|\phi(k)|^2$ are centered at $x = 0$ and $k = k_0$, respectively. It is convenient to define the half-widths Δx and Δk as corresponding to the half-maxima of $|\psi_0(x)|^2$ and $|\phi(k)|^2$. In this way, when x varies from 0 to $\pm\Delta x$ and k from k_0 to $k_0 \pm \Delta k$, the functions $|\psi_0(x)|^2$ and $|\phi(k)|^2$ drop to $e^{-1/2}$:

$$\frac{|\psi(\pm\Delta x, 0)|^2}{|\psi(0, 0)|^2} = e^{-1/2}, \qquad \frac{|\phi(k_0 \pm \Delta k)|^2}{|\phi(k_0)|^2} = e^{-1/2}. \tag{1.116}$$

These equations, combined with (1.115), lead to $e^{-2\Delta x^2/a^2} = e^{-1/2}$ and $e^{-a^2\Delta k^2/2} = e^{-1/2}$, respectively, or to

$$\Delta x = \frac{a}{2}, \qquad \Delta k = \frac{1}{a}; \tag{1.117}$$

hence

$$\Delta x \Delta k = \frac{1}{2}. \tag{1.118}$$

Since $\Delta k = \Delta p/\hbar$ we have

$$\boxed{\Delta x \Delta p = \frac{\hbar}{2}.} \tag{1.119}$$

This relation shows that if the packet's width is narrow in x-space, its width in momentum space must be very broad, and vice versa.

A comparison of (1.119) with Heisenberg's uncertainty relations (1.62) reveals that the Gaussian wave packet yields an *equality*, not an *inequality* relation. In fact, equation (1.119) is the *lowest limit* of Heisenberg's inequality. As a result, the Gaussian wave packet is called the *minimum uncertainty* wave packet. All other wave packets yield higher values for the product of the x and p uncertainties: $\Delta x \Delta p > \hbar/2$; for an illustration see Problem 1.11. In conclusion, the value of the uncertainties product $\Delta x \Delta p$ varies with the choice of ψ, but the lowest bound, $\hbar/2$, is provided by a Gaussian wave function. We have now seen how the wave packet concept offers a heuristic way of deriving Heisenberg's uncertainty relations; a more rigorous derivation is given in Chapter 2.

1.8.3 Motion of Wave Packets

How do wave packets evolve in time? The answer is important, for it gives an idea not only about the motion of a quantum particle in space but also about the connection between classical and quantum mechanics. Besides studying how wave packets propagate in space, we will also examine the conditions under which packets may or may not spread.

At issue here is, knowing the initial wave packet $\psi_0(x)$ or the amplitude $\phi(k)$, how do we find $\psi(x, t)$ at any later time t? This issue reduces to calculating the integral $\int \phi(k) e^{i(kx-\omega t)} dk$ in (1.100). To calculate this integral, we need to specify the angular frequency ω and the amplitude $\phi(k)$. We will see that the spreading or nonspreading of the packet is dictated by the form of the function $\omega(k)$.

1.8.3.1 Propagation of a Wave Packet without Distortion

The simplest form of the angular frequency ω is when it is *proportional* to the wave number k; this case corresponds to a *nondispersive* propagation. Since the constant of proportionality has the dimension of a velocity[56], which we denote by v_0 (i.e., $\omega = v_0 k$), the wave packet (1.100) becomes

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{ik(x-v_0 t)} dk. \tag{1.120}$$

This relation has the same structure as (1.101), which suggests that $\psi(x, t)$ is identical with $\psi_0(x - v_0 t)$:

$$\psi(x, t) = \psi_0(x - v_0 t); \tag{1.121}$$

the form of the wave packet at time t is identical with the initial form. Therefore, when ω is proportional to k, so that $\omega = v_0 k$, the wave packet travels to the right with constant velocity v_0 *without distortion*.

However, since we are interested in wave packets that describe particles, we need to consider the more general case of *dispersive* media which transmit harmonic waves of different

[56] For propagation of light in a vacuum this constant is equal to c, the speed of light.

frequencies at different velocities. This means that ω is a *function* of k: $\omega = \omega(k)$. The form of $\omega(k)$ is determined by the requirement that the wave packet $\psi(x,t)$ describes the particle. Assuming that the amplitude $\phi(k)$ peaks at $k = k_0$, then $\phi(k) = g(k-k_0)$ is appreciably different from zero only in a narrow range $\Delta k = k - k_0$, and we can Taylor expand $\omega(k)$ about k_0:

$$\begin{aligned}\omega(k) &= \omega(k_0) + (k-k_0)\left.\frac{d\omega(k)}{dk}\right|_{k=k_0} + \frac{1}{2}(k-k_0)^2\left.\frac{d^2\omega(k)}{dk^2}\right|_{k=k_0} + \cdots \\ &= \omega(k_0) + (k-k_0)v_g + (k-k_0)^2\alpha + \cdots \end{aligned} \tag{1.122}$$

where $v_g = \left.\frac{d\omega(k)}{dk}\right|_{k=k_0}$ and $\alpha = \frac{1}{2}\left.\frac{d^2\omega(k)}{dk^2}\right|_{k=k_0}$.

Now, to determine $\psi(x,t)$ we need simply to substitute (1.122) into (1.100) with $\phi(k) = g(k-k_0)$. This leads to

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}}e^{ik_0(x-v_{ph}t)}\int_{-\infty}^{+\infty} g(k-k_0)e^{i(k-k_0)(x-v_g t)}e^{-i(k-k_0)^2\alpha t+\cdots}dk \tag{1.123}$$

where[57]

$$\boxed{v_g = \frac{d\omega(k)}{dk}, \qquad v_{ph} = \frac{\omega(k)}{k};} \tag{1.124}$$

v_{ph} and v_g are respectively the *phase* velocity and the *group* velocity. The phase velocity denotes the velocity of propagation for the phase of a single harmonic wave, $e^{ik_0(x-v_{ph}t)}$, and the group velocity represents the velocity of motion for the group of waves that make up the packet. One should not confuse the phase velocity and the group velocity; in general they are different. Only when ω is proportional to k will they be equal, as can be inferred from (1.124).

Group and phase velocities

Let us take a short detour to explain the meanings of v_{ph} and v_g. As mentioned above, when we superimpose many waves of different amplitudes and frequencies, we can obtain a wave packet or pulse which travels at the group velocity v_g; the *individual* waves that constitute the packet, however, move with different speeds; each wave moves with its own phase velocity v_{ph}. Figure 1.16 gives a qualitative illustration: the group velocity represents the velocity with which the wave packet propagates as a *whole*, where the individual waves (located inside the packet's envelope) that add up to make the packet move with different phase velocities. As shown in Figure 1.16, the wave packet has an appreciable magnitude only over a small region and falls rapidly outside this region.

The difference between the group velocity and the phase velocity can be understood quantitatively by deriving a relationship between them. A differentiation of $\omega = kv_{ph}$ (see (1.124)) with respect to k yields $d\omega/dk = v_{ph} + k(dv_{ph}/dk)$, and since $k = 2\pi/\lambda$, we have $dv_{ph}/dk = (dv_{ph}/d\lambda)(d\lambda/dk) = -(2\pi/k^2)(dv_{ph}/d\lambda)$ or $k(dv_{ph}/dk) = -\lambda(dv_{ph}/d\lambda)$; combining these relations, we obtain

$$\boxed{v_g = \frac{d\omega}{dk} = v_{ph} + k\frac{dv_{ph}}{dk} = v_{ph} - \lambda\frac{dv_{ph}}{d\lambda},} \tag{1.125}$$

[57] In these equations we have omitted k_0 since they are valid for any choice of k_0.

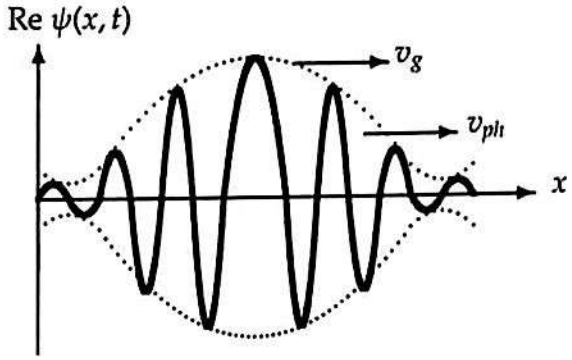


Figure 1.16: The function Re $\psi(x,t)$ of the wave packet (1.123), represented here by the solid curve contained in the dashed-curve envelope, propagates with the *group* velocity v_g along the x axis; the individual waves (not drawn here), which add up to make the solid curve, move with different *phase* velocities v_{ph}.

which we can also write as

$$\boxed{v_g = v_{ph} + p\frac{dv_{ph}}{dp},} \tag{1.126}$$

since $k(dv_{ph}/dk) = (p/\hbar)(dv_{ph}/dp)(dp/dk) = p(dv_{ph}/dp)$ because $k = p/\hbar$. Equations (1.125) and (1.126) show that the group velocity may be larger or smaller than the phase velocity; it may also be equal to the phase velocity depending on the medium. If the phase velocity does not depend on the wavelength—this occurs in nondispersive media—the group and phase velocities are equal, since $dv_{ph}/d\lambda = 0$. But if v_{ph} depends on the wavelength—this occurs in dispersive media—then $dv_{ph}/d\lambda \neq 0$; hence the group velocity may be smaller or larger than the phase velocity. An example of a nondispersive medium is an inextensible string; we would expect $v_g = v_{ph}$. Water waves offer a typical dispersive medium; in Problem 1.13 we show that for deepwater waves we have $v_g = \frac{1}{2}v_{ph}$ and for surface waves we have $v_g = \frac{3}{2}v_{ph}$; see (1.217) and (1.219).

Consider the case of a particle traveling in a *constant potential* V; its total energy is $E(p) = p^2/(2m)+V$. Since the corpuscular features (energy and momentum) of a particle are connected to its wave characteristics (wave frequency and number) by the relations $E = \hbar\omega$ and $p = \hbar k$, we can rewrite (1.124) as follows:

$$\boxed{v_g = \frac{dE(p)}{dp}, \qquad v_{ph} = \frac{E(p)}{p},} \tag{1.127}$$

which, when combined with $E(p) = \frac{p^2}{2m} + V$, yield

$$v_g = \frac{d}{dp}\left(\frac{p^2}{2m} + V\right) = \frac{p}{m} = v_{particle}, \qquad v_{ph} = \frac{1}{p}\left(\frac{p^2}{2m} + V\right) = \frac{p}{2m} + \frac{V}{p}. \tag{1.128}$$

The group velocity of the wave packet is thus equal to the classical velocity of the particle, $v_g = v_{particle}$. This suggests we should view the "center" of the wave packet as traveling like a classical particle that obeys the laws of classical mechanics: the center would then follow the "classical

trajectory" of the particle. We now see how the wave packet concept offers a clear connection between the classical description of a particle and its quantum mechanical description. In the case of a *free* particle, an insertion of $V = 0$ into (1.128) yields

$$v_g = \frac{p}{m}, \qquad v_{ph} = \frac{p}{2m} = \frac{1}{2}v_g. \tag{1.129}$$

This shows that, while the group velocity of the wave packet corresponding to a free particle is equal to the particle's velocity, p/m, the phase velocity is half the group velocity. The expression $v_{ph} = \frac{1}{2}v_g$ is meaningless, for it states that the wave function travels at half the speed of the particle it is intended to represent. This is unphysical indeed. The phase velocity has in general no meaningful physical significance.

Time-evolution of the packet

Having taken a short detour to discuss the phase and group velocities, let us now return to our main task of calculating the packet $\psi(x,t)$ as listed in (1.123). For this, we need to decide on where to terminate the expansion (1.122) or the exponent in the integrand of (1.123). We are going to consider two separate cases corresponding to whether we terminate the exponent in (1.123) at the linear term, $(k-k_0)v_g t$, or at the quadratic term, $(k-k_0)^2\alpha t$. These two cases are respectively known as the *linear approximation* and the *quadratic approximation*.

In the linear approximation, which is justified when $g(k-k_0)$ is narrow enough to neglect the quadratic k^2 term, $(k-k_0)^2\alpha t \ll 1$, the wave packet (1.123) becomes

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} e^{ik_0(x-v_{ph}t)} \int_{-\infty}^{+\infty} g(k-k_0) e^{i(k-k_0)(x-v_g t)} dk. \tag{1.130}$$

This relation can be rewritten as

$$\psi(x,t) = e^{ik_0(x-v_{ph}t)} \psi_0(x-v_g t) e^{-ik_0(x-v_g t)}, \tag{1.131}$$

where ψ_0 is the initial wave packet (see (1.101))

$$\psi_0(x-v_g t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(q) e^{i(x-v_g t)q + ik_0(x-v_g t)} dq; \tag{1.132}$$

the new variable q stands for $q = k - k_0$. Equation (1.131) leads to

$$\left|\psi(x,t)\right|^2 = \left|\psi_0(x-v_g t)\right|^2. \tag{1.133}$$

Equation (1.131) represents a wave packet whose amplitude is modulated. As depicted in Figure 1.16, the modulating wave, $\psi_0(x-v_g t)$, propagates to the right with the group velocity v_g; the modulated wave, $e^{ik_0(x-v_{ph}t)}$, represents a pure harmonic wave of constant wave number k_0 that also travels to the right with the phase velocity v_{ph}. That is, (1.131) and (1.133) represent a wave packet whose peak travels as a whole with the velocity v_g, while the individual wave propagates inside the envelope with the velocity v_{ph}. The group velocity, which gives the velocity of the packet's peak, clearly represents the velocity of the particle, since the chance of finding the particle around the packet's peak is much higher than finding it in any other region of space; the wave packet is highly localized in the neighborhood of the particle's position and vanishes elsewhere. It is therefore the group velocity, not the phase velocity, that is equal

to the velocity of the particle represented by the packet. This suggests that the motion of a material particle can be described well by wave packets. By establishing a correspondence between the particle's velocity and the velocity of the wave packet's peak, we see that the wave packet concept jointly embodies the particle aspect and the wave aspect of material particles.

Now, what about the size of the wave packet in the linear approximation? Is it affected by the particle's propagation? Clearly not. This can be inferred immediately from (1.131): $\psi_0(x - v_g t)$ represents, mathematically speaking, a curve that travels to the right with a velocity v_g without deformation. This means that if the packet is initially Gaussian, it will remain Gaussian as it propagates in space without any change in its size.

To summarize, we have shown that, in the linear approximation, the wave packet propagates undistorted and undergoes a uniform translational motion. Next we are going to study the conditions under which the packet experiences deformation.

1.8.3.2 Propagation of a Wave Packet with Distortion

Let us now include the quadratic k^2 term, $(k - k_0)^2 \alpha t$, in the integrand's exponent of (1.123) and drop the higher terms. This leads to

$$\psi(x,t) = e^{ik_0(x - v_{ph}t)} f(x,t), \tag{1.134}$$

where $f(x,t)$, which represents the envelope of the packet, is given by

$$f(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(q) e^{iq(x - v_g t)} e^{-iq^2 \alpha t} dq, \tag{1.135}$$

with $q = k - k_0$. Were it not for the quadratic q^2 correction, $iq^2\alpha t$, the wave packet would move uniformly without any change of shape, since similarly to (1.121), $f(x,t)$ would be given by $f(x,t) = \psi_0(x - v_g t)$.

To show how α affects the width of the packet, let us consider the Gaussian packet (1.108) whose amplitude is given by $\phi(k) = (a^2/2\pi)^{1/4} \exp\left[-a^2(k - k_0)^2/4\right]$ and whose initial width is $\Delta x_0 = a/2$ and $\Delta k = \hbar/a$. Substituting $\phi(k)$ into (1.134), we obtain

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{2\pi}\right)^{1/4} e^{ik_0(x - v_{ph}t)} \int_{-\infty}^{+\infty} \exp\left[iq(x - v_g t) - \left(\frac{a^2}{4} + i\alpha t\right) q^2\right] dq. \tag{1.136}$$

Evaluating the integral (the calculations are detailed in the following example, see Eq. (1.150)), we can show that the packet's density distribution is given by

$$\left|\psi(x,t)\right|^2 = \frac{1}{\sqrt{2\pi}\Delta x(t)} \exp\left\{-\frac{\left(x - v_g t\right)^2}{2\left[\Delta x(t)\right]^2}\right\}, \tag{1.137}$$

where $\Delta x(t)$ is the width of the packet at time t:

$$\Delta x(t) = \frac{a}{2}\sqrt{1 + \frac{16\alpha^2}{a^4} t^2} = \Delta x_0 \sqrt{1 + \frac{\alpha^2 t^2}{(\Delta x_0)^4}}. \tag{1.138}$$

We see that the packet's width, which was initially given by $\Delta x_0 = a/2$, has grown by a factor of $\sqrt{1 + a^2t^2/(\Delta x_0)^4}$ after time t. Hence the wave packet is spreading; *the spreading is due to the inclusion of the quadratic q^2 term, $iq^2\alpha t$*. Should we drop this term, the packet's width $\Delta x(t)$ would then remain constant, equal to Δx_0.

The density distribution (1.137) displays two results: (1) the center of the packet moves with the group velocity; (2) the packet's width increases linearly with time. From (1.138) we see that the packet begins to spread appreciably only when $a^2t^2/(\Delta x_0)^4 \approx 1$ or $t \approx (\Delta x_0)^2/\alpha$. In fact, if $t \ll (\Delta x_0)^2/\alpha$ the packet's spread will be negligible, whereas if $t \gg \frac{(\Delta x_0)^2}{\alpha}$ the packet's spread will be significant.

To be able to make concrete statements about the growth of the packet, as displayed in (1.138), we need to specify α; this reduces to determining the function $\omega(k)$, since $\alpha = \frac{1}{2}\frac{d^2\omega}{dk^2}\Big|_{k=k_0}$. For this, let us invoke an example that yields itself to explicit calculation. In fact, the example we are going to consider—a *free* particle with a Gaussian amplitude—allows the calculations to be performed *exactly*; hence there is no need to expand $\omega(k)$.

Example 1.9 (Free particle with a Gaussian wave packet)
Determine how the wave packet corresponding to a free particle, with an initial Gaussian packet, spreads in time.

Solution
The issue here is to find out how the wave packet corresponding to a free particle with $\phi(k) = (a^2/2\pi)^{1/4}e^{-a^2(k-k_0)^2/4}$ (see (1.115)) spreads in time.

First, we need to find the form of the wave packet, $\psi(x,t)$. Substituting the amplitude $\phi(k) = (a^2/2\pi)^{1/4}e^{-a^2(k-k_0)^2/4}$ into the Fourier integral (1.100), we obtain

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}}\left(\frac{a^2}{2\pi}\right)^{1/4}\int_{-\infty}^{+\infty}\exp\left[-\frac{a^2}{4}(k-k_0)^2 + i(kx-\omega t)\right]dk. \tag{1.139}$$

Since $\omega(k) = \hbar k^2/(2m)$ (the dispersion relation for a free particle), and using a change of variables $q = k - k_0$, we can write the exponent in the integrand of (1.139) as a perfect square for q:

$$\begin{aligned}
-\frac{a^2}{4}(k-k_0)^2 + i\left(kx - \frac{\hbar k^2}{2m}t\right) &= -\left(\frac{a^2}{4} + i\frac{\hbar t}{2m}\right)q^2 + i\left(x - \frac{\hbar k_0 t}{m}\right)q \\
&\quad + ik_0\left(x - \frac{\hbar k_0 t}{2m}\right) \\
&= -\alpha q^2 + i\left(x - \frac{\hbar k_0 t}{m}\right)q + ik_0\left(x - \frac{\hbar k_0 t}{2m}\right) \\
&= -\alpha\left[q - \frac{i}{2\alpha}\left(x - \frac{\hbar k_0 t}{m}\right)\right]^2 - \frac{1}{4\alpha}\left(x - \frac{\hbar k_0 t}{m}\right)^2 \\
&\quad + ik_0\left(x - \frac{\hbar k_0 t}{2m}\right),
\end{aligned} \tag{1.140}$$

where we have used the relation $-\alpha q^2 + iyq = -\alpha\left[q - iy/(2\alpha)\right]^2 - y^2/(4\alpha)$, with $y = x - \hbar k_0 t/m$ and

$$\alpha = \frac{a^2}{4} + i\frac{\hbar t}{2m}. \tag{1.141}$$

Substituting (1.140) into (1.139) we obtain

$$\begin{aligned}\psi(x,t) &= \frac{1}{\sqrt{2\pi}}\left(\frac{a^2}{2\pi}\right)^{1/4}\exp\left[ik_0\left(x - \frac{\hbar k_0 t}{2m}\right)\right]\exp\left[-\frac{1}{4\alpha}\left(x - \frac{\hbar k_0 t}{m}\right)^2\right] \\ &\quad\times\int_{-\infty}^{+\infty}\exp\left\{-\alpha\left[q - \frac{i}{2\alpha}\left(x - \frac{\hbar k_0 t}{m}\right)\right]^2\right\}dq.\end{aligned} \tag{1.142}$$

Combined with the integral[58] $\int_{-\infty}^{+\infty}\exp\left[-\alpha\left(q - iy/(2\alpha)\right)^2\right]dq = \sqrt{\pi/\alpha}$, (1.142) leads to

$$\psi(x,t) = \frac{1}{\sqrt{\alpha}}\left(\frac{a^2}{8\pi}\right)^{1/4}\exp\left[ik_0\left(x - \frac{\hbar k_0 t}{2m}\right)\right]\exp\left[-\frac{1}{4\alpha}\left(x - \frac{\hbar k_0 t}{m}\right)^2\right]. \tag{1.143}$$

Since α is a complex number (see (1.141)), we can write it in terms of its modulus and phase

$$\alpha = \frac{a^2}{4}\left(1 + i\frac{2\hbar t}{ma^2}\right) = \frac{a^2}{4}\left(1 + \frac{4\hbar^2 t^2}{m^2 a^4}\right)^{1/2} e^{i\theta}, \tag{1.144}$$

where $\theta = \tan^{-1}\left[2\hbar t/(ma^2)\right]$; hence

$$\frac{1}{\sqrt{\alpha}} = \frac{2}{a}\left(1 + \frac{4\hbar^2 t^2}{m^2 a^4}\right)^{-1/4} e^{-i\theta/2}. \tag{1.145}$$

Substituting (1.141) and (1.145) into (1.143), we have

$$\psi(x,t) = \left(\frac{2}{\pi a^2}\right)^{1/4}\left(1 + \frac{4\hbar^2 t^2}{m^2 a^4}\right)^{-1/4} e^{-i\theta/2} e^{ik_0(x - \hbar k_0 t/2m)}\exp\left[-\frac{(x - \hbar k_0 t/m)^2}{a^2 + 2i\hbar t/m}\right]. \tag{1.146}$$

Since $\left|e^{-y^2/(a^2+2i\hbar t/m)}\right|^2 = e^{-y^2/(a^2-2i\hbar t/m)} e^{-y^2/(a^2+2i\hbar t/m)}$, where $y = x - \hbar k_0 t/m$, and since $y^2/(a^2 - 2i\hbar t/m) + y^2/(a^2 + 2i\hbar t/m) = 2a^2 y^2/(a^4 + 4\hbar^2 t^2/m^2)$, we have

$$\left|\exp\left(-\frac{y^2}{a^2 + 2i\hbar t/m}\right)\right|^2 = \exp\left(-\frac{2a^2 y^2}{a^4 + 4\hbar^2 t^2/m^2}\right); \tag{1.147}$$

hence

$$\begin{aligned}\left|\psi(x,t)\right|^2 &= \sqrt{\frac{2}{\pi a^2}}\left(1 + \frac{4\hbar^2 t^2}{m^2 a^4}\right)^{-1/2}\left|\exp\left[-\frac{(x - \hbar k_0 t/m)^2}{a^2 + 2i\hbar t/m}\right]\right|^2 \\ &= \sqrt{\frac{2}{\pi a^2}}\frac{1}{\gamma(t)}\exp\left\{-\frac{2}{\left[a\gamma(t)\right]^2}\left(x - \frac{\hbar k_0 t}{m}\right)^2\right\},\end{aligned} \tag{1.148}$$

[58] If β and δ are two complex numbers and if $\mathrm{Re}\,\beta > 0$, we have $\int_{-\infty}^{+\infty} e^{-\beta(q+\delta)^2}dq = \sqrt{\pi/\beta}$.

where $\gamma(t) = \sqrt{1 + 4\hbar^2 t^2/(m^2 a^4)}$.

We see that both the wave packet (1.146) and the probability density (1.148) remain Gaussian as time evolves. This can be traced to the fact that the x-dependence of the phase, $e^{ik_0 x}$, of $\psi_0(x)$ as displayed in (1.115) is *linear*. If the x-dependence of the phase were other than linear, say quadratic, the form of the wave packet would not remain Gaussian. So the phase factor $e^{ik_0 x}$, which was present in $\psi_0(x)$, allows us to account for the motion of the particle.

Since the group velocity of a free particle is $v_g = d\omega/dk = \frac{d}{dk}\left(\frac{\hbar k^2}{2m}\right)\Big|_{k_0} = \hbar k_0/m$, we can rewrite (1.146) as follows[59]:

$$\psi(x,t) = \frac{1}{\sqrt{\sqrt{2\pi}\Delta x(t)}} e^{-i\theta/2} e^{ik_0(x - v_g t/2)} \exp\left[-\frac{(x - v_g t)^2}{a^2 + 2i\hbar t/m}\right], \tag{1.149}$$

$$\left|\psi(x,t)\right|^2 = \frac{1}{\sqrt{2\pi}\Delta x(t)} \exp\left\{-\frac{(x - v_g t)^2}{2\left[\Delta x(t)\right]^2}\right\}, \tag{1.150}$$

where[60]

$$\Delta x(t) = \frac{a}{2}\gamma(t) = \frac{a}{2}\sqrt{1 + \frac{4\hbar^2 t^2}{m^2 a^4}} \tag{1.151}$$

represents the width of the wave packet at time t. Equations (1.149) and (1.150) describe a Gaussian wave packet that is centered at $x = v_g t$ whose peak travels with the group speed $v_g = \hbar k_0/m$ and whose width $\Delta x(t)$ increases linearly with time. So, during time t, the packet's center has moved from $x = 0$ to $x = v_g t$ and its width has expanded from $\Delta x_0 = a/2$ to $\Delta x(t) = \Delta x_0 \sqrt{1 + 4\hbar^2 t^2/(m^2 a^4)}$. The wave packet therefore undergoes a distortion; although it remains Gaussian, its width broadens linearly with time whereas its height, $1/(\sqrt{2\pi}\Delta x(t))$, decreases with time. As depicted in Figure 1.17, the wave packet, which had a very broad width and a very small amplitude at $t \to -\infty$, becomes narrower and narrower and its amplitude larger and larger as time increases towards $t = 0$; at $t = 0$ the packet is very localized, its width and amplitude being given by $\Delta x_0 = a/2$ and $\sqrt{2/\pi a^2}$, respectively. Then, as time increases ($t > 0$), the width of the packet becomes broader and broader, and its amplitude becomes smaller and smaller.

In the rest of this section we are going to comment on several features that are relevant not only to the Gaussian packet considered above but also to more general wave packets. First, let us begin by estimating the time at which the wave packet starts to spread out appreciably. The packet, which is initially narrow, begins to grow out noticeably only when the second term, $2\hbar t/(ma^2)$, under the square root sign of (1.151) is of order unity. For convenience, let us

[59] It is interesting to note that the harmonic wave $e^{ik_0(x - v_g t/2)}$ propagates with a phase velocity which is *half* the group velocity; as shown in (1.129), this is a property of *free* particles.

[60] We can derive (1.151) also from (1.116): a combination of the half-width $\left|\psi(\pm\Delta x, t)\right|^2 / \left|\psi(0,0)\right|^2 = e^{-1/2}$ with (1.148) yields $e^{-2\left[\Delta x/a\gamma(t)\right]^2} = e^{-1/2}$, which in turn leads to (1.151).

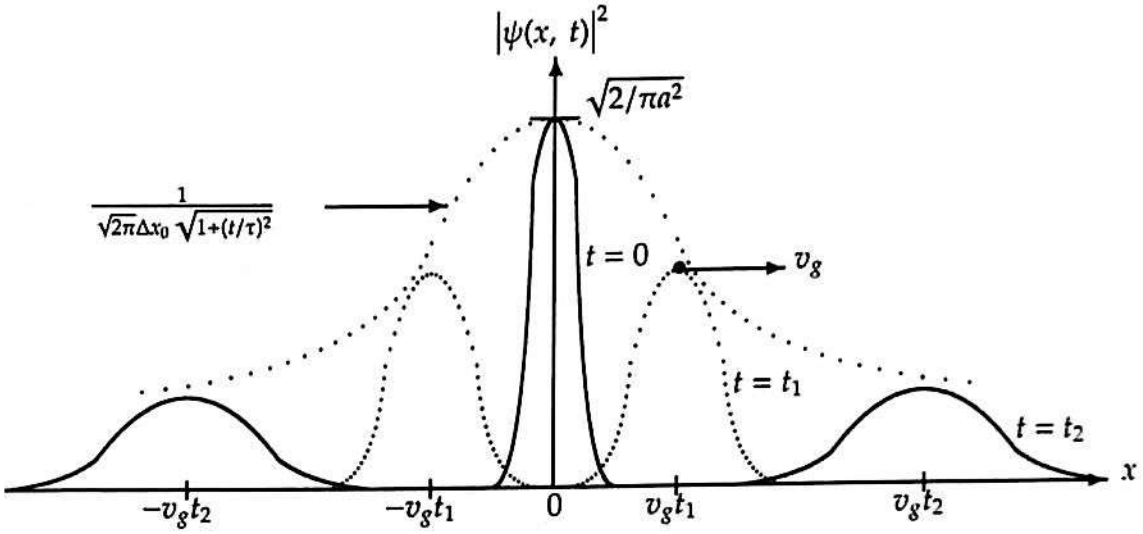


Figure 1.17: Time evolution of $|\psi(x, t)|^2$: the peak of the packet, which is centered at $x = v_g t$, moves with the speed v_g from left to right. The height of the packet, represented here by the dotted envelope, is modulated by the function $1/(\sqrt{2\pi}\Delta x(t))$, which goes to zero at $t \to \pm\infty$ and is equal to $\sqrt{2/\pi a^2}$ at $t = 0$. The width of the packet $\Delta x(t) = \Delta x_0\sqrt{1 + (t/\tau)^2}$ increases linearly with time.

write (1.151) in the form

$$\Delta x(t) = \Delta x_0\sqrt{1 + \left(\frac{t}{\tau}\right)^2}, \tag{1.152}$$

where

$$\tau = \frac{2m(\Delta x_0)^2}{\hbar} \tag{1.153}$$

represents a time constant that characterizes the rate of the packet's spreading. Now we can estimate the order of magnitude of τ; it is instructive to evaluate it for *microscopic* particles as well as for *macroscopic* particles. For instance, τ for an electron whose position is defined to within 10^{-10} m is given by[61] $\tau \simeq 1.7 \times 10^{-16}$ s; on the other hand, the time constant for a macroscopic particle of mass say 1 g whose position is defined to within 1 mm is of the order[62] of $\tau \simeq 2\times 10^{25}$ s (for an illustration see Problems 1.15 and 1.16). This crude calculation suggests that the wave packets of microscopic systems very quickly undergo significant growth; as for the packets of macroscopic systems, they begin to grow out noticeably only after the system has been in motion for an absurdly long time, a time of the order of, if not much higher than, the age of the Universe itself, which is about 4.7×10^{17} s. Having estimated the times at which the packet's spread becomes appreciable, let us now shed some light on the size of the spread. From (1.152) we see that when $t \gg \tau$ the packet's spreading is significant and, conversely, when $t \ll \tau$ the spread is negligible. As the cases $t \gg \tau$ and $t \ll \tau$ correspond to microscopic and macroscopic systems, respectively, we infer that the packet's dispersion is significant for microphysical systems and negligible for macroscopic systems. In the case of macroscopic

[61] If $\Delta x_0 = 10^{-10}$ m and since the rest mass energy of an electron is $mc^2 = 0.5$ MeV and using $\hbar c \simeq 197\times 10^{-15}$ MeV m, we have $\tau = 2mc^2(\Delta x_0)^2/((\hbar c)c) \simeq 1.7 \times 10^{-16}$ s.

[62] Since $\hbar = 1.05 \times 10^{-34}$ J s we have $\tau = 2 \times 0.001$ kg $\times$ (0.001 m)^2/(1.05 $\times 10^{-34}$ J s) $\simeq 2 \times 10^{25}$ s.

systems, the spread is there but it is too small to detect. For an illustration see Problem 1.15 where we show that the width of a 100 g object increases by an absurdly small factor of about 10^{-29} after traveling a distance of 100 m, but the width of a 25 eV electron increases by a factor of 10^9 after traveling the same distance (in a time of 3.3×10^{-5} s). Such an immense dispersion in such a short time is indeed hard to visualize classically; this motion cannot be explained by classical physics.

So the wave packets of propagating, microscopic particles are prone to spreading out very significantly in a short time. This spatial spreading seems to generate a conceptual problem: the spreading is incompatible with our expectation that the packet should remain highly localized at all times. After all, the wave packet is supposed to represent the particle and, as such, it is expected to travel without dispersion. For instance, the charge of an electron does not spread out while moving in space; the charge should remain localized inside the corresponding wave packet. In fact, whenever microscopic particles (electrons, neutrons, protons, etc.) are observed, they are always confined to small, finite regions of space; they never spread out as suggested by equation (1.151). How do we explain this apparent contradiction? The problem here has to do with the proper interpretation of the situation: we must modify the classical concepts pertaining to the meaning of the position of a particle. The wave function (1.146) cannot be identified with a material particle. The quantity $|\psi(x, t)|^2 dx$ represents the probability (Born's interpretation) of finding the particle described by the packet $\psi(x, t)$ at time t in the spatial region located between x and $x + dx$. *The material particle does not disperse (or fuzz out); yet its position cannot be known exactly.* The spreading of the matter wave, which is accompanied by a shrinkage of its height, as indicated in Figure 1.17, corresponds to a decrease of the probability density $|\psi(x, t)|^2$ and implies in no way a growth in the size of the particle. So the wave packet gives only the probability that the particle it represents will be found at a given position. No matter how broad the packet becomes, we can show that its norm is always conserved, for it does not depend on time. In fact, as can be inferred from (1.148), the norm of the packet is equal to one:

$$\int_{-\infty}^{+\infty} |\psi(x,t)|^2 \, dx = \sqrt{\frac{2}{\pi a^2}} \frac{1}{\gamma} \int_{-\infty}^{+\infty} \exp\left\{ -\frac{2\,(x - \hbar k_0 t/m)^2}{(a\gamma)^2} \right\} dx = \sqrt{\frac{2}{\pi a^2}} \frac{1}{\gamma} \sqrt{\frac{\pi a^2 \gamma^2}{2}} = 1, \tag{1.154}$$

since $\int_{-\infty}^{+\infty} e^{-ax^2} dx = \sqrt{\pi/a}$. This is expected, since the probability of finding the particle somewhere along the x-axis must be equal to one. The important issue here is that the norm of the packet is time independent and that its spread does not imply that the material particle becomes bloated during its motion, but simply implies a redistribution of the probability density. So, in spite of the significant spread of the packets of microscopic particles, the norms of these packets are always conserved—normalized to unity.

Besides, we should note that the example considered here is an *idealized* case, for we are dealing with a *free* particle. If the particle is subject to a potential, as in the general case, its wave packet will not spread as dramatically as that of a free particle. In fact, a varying potential can cause the wave packet to become narrow. This is indeed what happens when a measurement is performed on a microscopic system; the interaction of the system with the measuring device makes the packet very narrow, as will be seen in Chapter 3.

Let us now study how the spreading of the wave packet affects the uncertainties product $\Delta x(t)\Delta p(t)$. First, we should point out that the average momentum of the packet $\hbar k_0$ and its uncertainty $\hbar \Delta k$ do not change in time. This can be easily inferred as follows. Rewriting

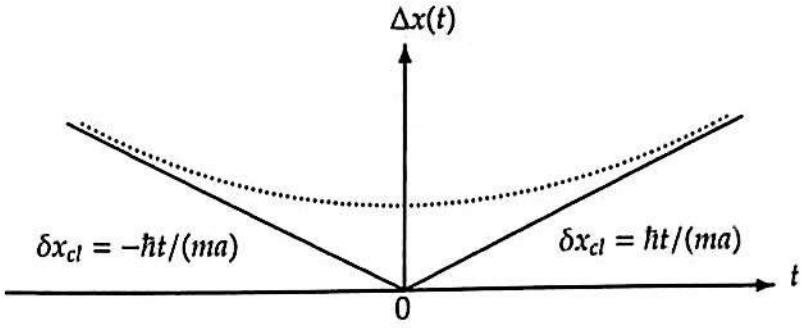


Figure 1.18: Time evolutions of the packet's width $\Delta x(t) = \Delta x_0 \sqrt{1 + (\delta x_{cl}(t)/\Delta x_0)^2}$ (dotted curve) and of the classical dispersion $\delta x_{cl}(t) = \pm\hbar t/(ma)$ (solid lines). For large values of $|t|$, $\Delta x(t)$ approaches $\delta x_{cl}(t)$ and at $t = 0$, $\Delta x(0) = \Delta x_0 = a/2$.

(1.100) in the form

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k,0) e^{i(kx-\omega t)} dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k,t) e^{ikx} dk, \tag{1.155}$$

we have

$$\phi(k,t) = e^{-i\omega(k)t} \phi(k,0), \tag{1.156}$$

where $\phi(k,0) = (a^2/2\pi)^{1/4} e^{-a^2(k-k_0)^2/4}$; hence

$$\left|\phi(k,t)\right|^2 = \left|\phi(k,0)\right|^2. \tag{1.157}$$

This suggests that the widths of $\phi(k,t)$ and $\phi(k,0)$ are equal; hence Δk remains constant and so must the momentum dispersion Δp (this is expected because the momentum of a free particle is a constant of the motion). Since the width of $\phi(k,0)$ is given by $\Delta k = 1/a$ (see (1.117)), we have

$$\Delta p = \hbar \Delta k = \frac{\hbar}{a}. \tag{1.158}$$

Multiplying this relation by (1.151), we have

$$\Delta x(t) \Delta p = \frac{\hbar}{2} \sqrt{1 + \frac{4\hbar^2}{m^2 a^4} t^2}, \tag{1.159}$$

which shows that $\Delta x(t)\Delta p \geq \hbar/2$ is satisfied at all times. Notably, when $t = 0$ we obtain the lower bound limit $\Delta x_0 \Delta p = \hbar/2$; this is the uncertainty relation for a stationary Gaussian packet (see (1.119)). As $|t|$ increases, however, we obtain an inequality, $\Delta x(t)\Delta p > \hbar/2$.

Having shown that the width of the packet does not disperse in momentum space, let us now study the dispersion of the packet's width in x-space. Since $\Delta x_0 = a/2$ we can write (1.151) as

$$\Delta x(t) = \frac{a}{2} \sqrt{1 + \frac{4\hbar^2 t^2}{m^2 a^4}} = \Delta x_0 \sqrt{1 + \left(\frac{\delta x_{cl}(t)}{\Delta x_0}\right)^2}, \tag{1.160}$$

where the dispersion factor $\delta x_{cl}(t)/\Delta x_0$ is given by

$$\frac{\delta x_{cl}(t)}{\Delta x_0} = \pm \frac{2\hbar}{ma^2} t = \pm \frac{\hbar}{2m\Delta x_0^2} t; \tag{1.161}$$

As shown in Figure 1.18, when $|t|$ is large (i.e., $t \to \pm\infty$), we have $\Delta x(t) \to \delta x_{cl}(t)$ with

$$\delta x_{cl}(t) = \pm\frac{\hbar t}{ma} = \pm\frac{\Delta p}{m}t = \pm\Delta v t, \tag{1.162}$$

where $\Delta v = \hbar/(ma)$ represents the dispersion in velocity. This means that if a particle starts initially ($t = 0$) at $x = 0$ with a velocity dispersion equal to Δv, then Δv will remain constant but the dispersion of the particle's position will increase linearly with time: $\delta x_{cl}(t) = \hbar|t|/(ma)$ (Figure 1.18). We see from (1.160) that if $\delta x_{cl}(t)/\Delta x_0 \ll 1$, the spreading of the wave packet is negligible, but if $\delta x_{cl}(t)/\Delta x_0 \gg 1$, the wave packet will spread out without bound.

We should highlight at this level the importance of the classical limit of (1.159): in the limit $\hbar \to 0$, the product $\Delta x(t)\Delta p$ goes to zero. This means that the x and p uncertainties become negligible; that is, in the classical limit, the wave packet will propagate without spreading. In this case the center of the wave packet moves like a free particle that obeys the laws of classical mechanics. *The spread of wave packets is thus a purely quantum effect.* So when $\hbar \to 0$ all quantum effects, the spread of the packet, disappear.

We may conclude this study of wave packets by highlighting their importance:

- They provide a linkage with the Heisenberg uncertainty principle.
- They embody and unify the particle and wave features of matter waves.
- They provide a linkage between wave intensities and probabilities.
- They provide a connection between classical and quantum mechanics.

1.9 Concluding Remarks

Despite its striking success in predicting the hydrogen's energy levels and transition rates, the Bohr model suffers from a number of limitations:

- It works only for hydrogen and hydrogen-like ions such as $He^+$and Li^{2+}.
- It provides no explanation for the origin of its various assumptions. For instance, it gives no theoretical justification for the quantization condition (1.68) nor does it explain why stationary states radiate no energy.
- It fails to explain why, instead of moving *continuously* from one energy level to another, the electrons *jump* from one level to the other.

The model therefore requires considerable extension to account for the electronic properties and spectra of a wide range of atoms. Even in its present limited form, Bohr's model represents a bold and major departure from classical physics: classical physics offers no justification for the existence of discrete energy states in a system such as a hydrogen atom and no justification for the quantization of the angular momentum.

In its present form, the model not only suffers from incompleteness but also lacks the ingredients of a consistent theory. It was built upon a series of ad hoc, piecemeal assumptions. These assumptions were not derived from the first principles of a more general theory, but postulated rather arbitrarily.

The formulation of the theory of quantum mechanics was largely precipitated by the need to find a theoretical foundation for Bohr's ideas as well as to explain, from first principles, a wide variety of other microphysical phenomena such as the puzzling processes discussed in this chapter. It is indeed surprising that a single theory, quantum mechanics, is powerful and rich enough to explain accurately a wide variety of phenomena taking place at the molecular, atomic, and subatomic levels.

In this chapter we have dealt with the most important experimental facts which confirmed the failure of classical physics and subsequently led to the birth of quantum mechanics. In the rest of this text we will focus on the formalism of quantum mechanics and on its application to various microphysical processes. To prepare for this task, we need first to study the mathematical tools necessary for understanding the formalism of quantum mechanics; this is taken up in Chapter 2.

1.10 Solved Problems

Numerical calculations in quantum physics can be made simpler by using the following units. First, it is convenient to express energies in units of electronvolt (eV): one eV is defined as the energy acquired by an electron passing through a potential difference of one Volt. The electronvolt unit can be expressed in terms of joules and vice versa: 1 eV = $(1.6 \times 10^{-19}\ \text{C}) \times (1\ \text{V}) = 1.6 \times 10^{-19}\ \text{J}$ and $1\ \text{J} = 0.625 \times 10^{19}$ eV.

It is also convenient to express the masses of subatomic particles, such as the electron, proton, and neutron, in terms of their rest mass energies: $m_e c^2 = 0.511$ MeV, $m_p c^2 = 938.27$ MeV, and $m_n c^2 = 939.56$ MeV.

In addition, the quantities $\hbar c = 197.33$ MeV fm $= 197.33 \times 10^{-15}$ MeV m or $hc = 1242.37 \times 10^{-10}$ eV m are sometimes more convenient to use than $\hbar = 1.05 \times 10^{-34}$ J s. Additionally, instead of $1/(4\pi\varepsilon_0) = 8.9 \times 10^9\ \text{N m}^2\ \text{C}^{-2}$, one should sometimes use the fine structure constant $\alpha = e^2/[(4\pi\varepsilon_0)\hbar c] = 1/137$.

Problem 1.1

A 45 kW broadcasting antenna emits radio waves at a frequency of 4 MHz.

(a) How many photons are emitted per second?

(b) Is the quantum nature of the electromagnetic radiation important in analyzing the radiation emitted from this antenna?

Solution

(a) The electromagnetic energy emitted by the antenna in one second is $E = 45\,000$ J. Thus, the number of photons emitted in one second is

$$n = \frac{E}{h\nu} = \frac{45\,000\ \text{J}}{6.63 \times 10^{-34}\ \text{J s} \times 4 \times 10^6\ \text{Hz}} = 1.7 \times 10^{31}. \tag{1.163}$$

(b) Since the antenna emits a huge number of photons every second, 1.7×10^{31}, the quantum nature of this radiation is unimportant. As a result, this radiation can be treated fairly accurately by the classical theory of electromagnetism.

Problem 1.2
Consider a mass–spring system where a 4 kg mass is attached to a massless spring of constant $k = 196\ \mathrm{N\,m^{-1}}$; the system is set to oscillate on a frictionless, horizontal table. The mass is pulled 25 cm away from the equilibrium position and then released.

(a) Use classical mechanics to find the total energy and frequency of oscillations of the system.

(b) Treating the oscillator with quantum theory, find the energy spacing between two consecutive energy levels and the total number of quanta involved. Are the quantum effects important in this system?

Solution

(a) According to classical mechanics, the frequency and the total energy of oscillations are given by

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{m}} = \frac{1}{2\pi}\sqrt{\frac{196}{4}} = 1.11\ \mathrm{Hz}, \qquad E = \frac{1}{2}kA^2 = \frac{196}{2}(0.25)^2 = 6.125\ \mathrm{J}. \tag{1.164}$$

(b) The energy spacing between two consecutive energy levels is given by

$$\Delta E = h\nu = (6.63 \times 10^{-34}\ \mathrm{J\ s}) \times (1.11\ \mathrm{Hz}) = 7.4 \times 10^{-34}\ \mathrm{J} \tag{1.165}$$

and the total number of quanta is given by

$$n = \frac{E}{\Delta E} = \frac{6.125\ \mathrm{J}}{7.4 \times 10^{-34}\ \mathrm{J}} = 8.3 \times 10^{33}. \tag{1.166}$$

We see that the energy of one quantum, 7.4×10^{-34} J, is completely negligible compared to the total energy 6.125 J, and that the number of quanta is very large. As a result, the energy levels of the oscillator can be viewed as continuous, for it is not feasible classically to measure the spacings between them. Although the quantum effects are present in the system, they are beyond human detection. So quantum effects are negligible for macroscopic systems.

Problem 1.3
When light of a given wavelength is incident on a metallic surface, the stopping potential for the photoelectrons is 3.2 V. If a second light source whose wavelength is double that of the first is used, the stopping potential drops to 0.8 V. From these data, calculate

(a) the wavelength of the first radiation and

(b) the work function and the cutoff frequency of the metal.

Solution

(a) Using (1.28) and since the wavelength of the second radiation is double that of the first one, $\lambda_2 = 2\lambda_1$, we can write

$$V_{s_1} = \frac{hc}{e\lambda_1} - \frac{W}{e}, \tag{1.167}$$

$$V_{s_2} = \frac{hc}{e\lambda_2} - \frac{W}{e} = \frac{hc}{2e\lambda_1} - \frac{W}{e}. \tag{1.168}$$

To obtain λ_1 we have only to subtract (1.168) from (1.167):

$$V_{s_1} - V_{s_2} = \frac{hc}{e\lambda_1}\left(1 - \frac{1}{2}\right) = \frac{hc}{2e\lambda_1}. \tag{1.169}$$

The wavelength is thus given by

$$\lambda_1 = \frac{hc}{2e(V_{s_1} - V_{s_2})} = \frac{6.6 \times 10^{-34}\ \text{J s} \times 3 \times 10^{8}\ \text{m s}^{-1}}{2 \times 1.6 \times 10^{-19}\ \text{C} \times (3.2\ \text{V} - 0.8\ \text{V})} = 2.6 \times 10^{-7}\ \text{m}. \tag{1.170}$$

(b) To obtain the work function, we simply need to multiply (1.168) by 2 and subtract the result from (1.167), $V_{s_1} - 2V_{s_2} = W/e$, which leads to

$$W = e(V_{s_1} - 2V_{s_2}) = 1.6\ \text{eV} = 1.6 \times 1.6 \times 10^{-19} = 2.56 \times 10^{-19}\ \text{J}. \tag{1.171}$$

The cutoff frequency is

$$\nu = \frac{W}{h} = \frac{2.56 \times 10^{-19}\ \text{J}}{6.6 \times 10^{-34}\ \text{J s}} = 3.9 \times 10^{14}\ \text{Hz}. \tag{1.172}$$

Problem 1.4

(a) Estimate the energy of the electrons that we need to use in an electron microscope to resolve a separation of 0.27 nm.

(b) In a scattering of 80 eV electrons from a crystal, the fifth maximum of the intensity is observed at an angle of 30°. Estimate the crystal's planar separation.

Solution

(a) Since the electron's momentum is $p = 2\pi\hbar/\lambda$, its kinetic energy is given by

$$E = \frac{p^2}{2m_e} = \frac{2\pi^2\hbar^2}{m_e\lambda^2}. \tag{1.173}$$

Since $m_e c^2 = 0.511$ MeV, $\hbar c = 197.33 \times 10^{-15}$ MeV m, and $\lambda = 0.27 \times 10^{-9}$ m, we have

$$E = \frac{2\pi^2(\hbar c)^2}{(m_e c^2)\lambda^2} = \frac{2\pi^2(197.33 \times 10^{-15}\ \text{MeV m})^2}{(0.511\ \text{MeV})(0.27 \times 10^{-9}\ \text{m})^2} = 20.6\ \text{eV}. \tag{1.174}$$

(b) Using Bragg's relation (1.51), $\lambda = (2d/n)\sin\phi$, where d is the crystal's planar separation, we can infer the electron's kinetic energy from (1.173):

$$E = \frac{p^2}{2m_e} = \frac{2\pi^2\hbar^2}{m_e\lambda^2} = \frac{n^2\pi^2\hbar^2}{2m_e d^2 \sin^2\phi}, \tag{1.175}$$

which leads to

$$d = \frac{n\pi\hbar}{(\sin\phi)\sqrt{2m_e E}} = \frac{n\pi\hbar c}{(\sin\phi)\sqrt{2m_e c^2 E}}. \tag{1.176}$$

Since $n = 5$ (the fifth maximum), $\phi = 30°$, $E = 2$ eV, and $m_e c^2 = 0.511$ MeV, we have

$$d = \frac{5\pi \times 197.33 \times 10^{-15}\ \text{MeV m}}{(\sin 30°)\sqrt{2 \times 0.511\ \text{MeV} \times 80 \times 10^{-6}\ \text{MeV}}} = 0.686\ \text{nm}. \tag{1.177}$$

Problem 1.5
A photon of energy 3 keV collides elastically with an electron initially at rest. If the photon emerges at an angle of 60°, calculate
(a) the kinetic energy of the recoiling electron and
(b) the angle at which the electron recoils.

Solution
(a) From energy conservation, we have

$$h\nu + m_e c^2 = h\nu' + (K_e + m_e c^2), \tag{1.178}$$

where $h\nu$ and $h\nu'$ are the energies of the initial and scattered photons, respectively, $m_e c^2$ is the rest mass energy of the initial electron, $(K_e + m_e c^2)$ is the total energy of the recoiling electron, and K_e is its recoil kinetic energy. The expression for K_e can immediately be inferred from (1.178):

$$K_e = h(\nu - \nu') = hc\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right) = \frac{hc}{\lambda}\frac{\lambda' - \lambda}{\lambda'} = (h\nu)\frac{\Delta\lambda}{\lambda'}, \tag{1.179}$$

where the wave shift $\Delta\lambda$ is given by (1.41):

$$\begin{aligned} \Delta\lambda &= \lambda' - \lambda = \frac{h}{m_e c}(1 - \cos\theta) = \frac{2\pi\hbar c}{m_e c^2}(1 - \cos\theta) \\ &= \frac{2\pi \times 197.33 \times 10^{-15}\ \text{MeV m}}{0.511\ \text{MeV}}(1 - \cos 60°) \\ &= 0.0012\ \text{nm}. \end{aligned} \tag{1.180}$$

Since the wavelength of the incident photon is $\lambda = 2\pi\hbar c/(h\nu)$, we have $\lambda = 2\pi \times 197.33 \times 10^{-15}$ MeV m/(0.003 MeV) = 0.414 nm; the wavelength of the scattered photon is given by

$$\lambda' = \lambda + \Delta\lambda = 0.4152\ \text{nm}. \tag{1.181}$$

Now, substituting the numerical values of λ' and $\Delta\lambda$ into (1.179), we obtain the kinetic energy of the recoiling electron

$$K_e = (h\nu)\frac{\Delta\lambda}{\lambda'} = (3\ \text{keV}) \times \frac{0.0012\ \text{nm}}{0.4152\ \text{nm}} = 8.671\ \text{eV}. \tag{1.182}$$

(b) To obtain the angle at which the electron recoils, we need simply to use the conservation of the total momentum along the $x-$ and $y-$ axes:

$$p = p_e \cos\phi + p' \cos\theta, \qquad 0 = p_e \sin\phi - p' \sin\theta. \tag{1.183}$$

These can be rewritten as

$$p_e \cos\phi = p - p' \cos\theta, \qquad p_e \sin\phi = p' \sin\theta, \tag{1.184}$$

where p and p' are the momenta of the initial and final photons, p_e is the momentum of the recoiling electron, and θ and ϕ are the angles at which the photon and electron scatter, respectively (Figure 1.5). Taking (1.184) and dividing the second equation by the first, we obtain

$$\tan\phi = \frac{\sin\theta}{p/p' - \cos\theta} = \frac{\sin\theta}{\lambda'/\lambda - \cos\theta}, \tag{1.185}$$

where we have used the momentum expressions of the incident photon $p = h/\lambda$ and of the scattered photon $p' = h/\lambda'$. Since $\lambda = 0.414$ nm and $\lambda' = 0.4152$ nm, the angle at which the electron recoils is given by

$$\phi = \tan^{-1}\left(\frac{\sin\theta}{\lambda'/\lambda - \cos\theta}\right) = \tan^{-1}\left(\frac{\sin 60^\circ}{0.4152/0.414 - \cos 60^\circ}\right) = 59.86^\circ. \tag{1.186}$$

Problem 1.6
Show that the maximum kinetic energy transferred to a proton when hit by a photon of energy $h\nu$ is $K_p = h\nu/[1 + m_p c^2/(2h\nu)]$, where m_p is the mass of the proton.

Solution
Using (1.40), we have

$$\frac{1}{\nu'} = \frac{1}{\nu} + \frac{h}{m_p c^2}(1 - \cos\theta), \tag{1.187}$$

which leads to

$$h\nu' = \frac{h\nu}{1 + (h\nu/m_p c^2)(1 - \cos\theta)}. \tag{1.188}$$

Since the kinetic energy transferred to the proton is given by $K_p = h\nu - h\nu'$, we obtain

$$K_p = h\nu - \frac{h\nu}{1 + (h\nu/m_p c^2)(1 - \cos\theta)} = \frac{h\nu}{1 + m_p c^2/[h\nu(1 - \cos\theta)]}. \tag{1.189}$$

Clearly, the maximum kinetic energy of the proton corresponds to the case where the photon scatters backwards ($\theta = \pi$),

$$K_p = \frac{h\nu}{1 + m_p c^2/(2h\nu)}. \tag{1.190}$$

Problem 1.7
Consider a photon that scatters from an electron at rest. If the Compton wavelength shift is observed to be triple the wavelength of the incident photon and if the photon scatters at 60°, calculate

(a) the wavelength of the incident photon,
(b) the energy of the recoiling electron, and
(c) the angle at which the electron scatters.

Solution
(a) In the case where the photons scatter at $\theta = 60^\circ$ and since $\Delta\lambda = 3\lambda$, the wave shift relation (1.41) yields

$$3\lambda = \frac{h}{m_e c}(1 - \cos 60^\circ), \tag{1.191}$$

which in turn leads to

$$\lambda = \frac{h}{6m_e c} = \frac{\pi\hbar c}{3m_e c^2} = \frac{3.14 \times 197.33 \times 10^{-15}\ \text{MeV m}}{3 \times 0.511\ \text{MeV}} = 4.04 \times 10^{-13}\ \text{m}. \tag{1.192}$$

(b) The energy of the recoiling electron can be obtained from the conservation of energy:

$$K_e = hc\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right) = \frac{3hc}{4\lambda} = \frac{3\pi\hbar c}{2\lambda} = \frac{3 \times 3.14 \times 197.33 \times 10^{-15}\ \text{MeV m}}{2 \times 4.04 \times 10^{-13}\ \text{m}} = 2.3\ \text{MeV}. \quad (1.193)$$

In deriving this relation, we have used the fact that $\lambda' = \lambda + \Delta\lambda = 4\lambda$.

(c) Since $\lambda' = 4\lambda$ the angle ϕ at which the electron recoils can be inferred from (1.186)

$$\phi = \tan^{-1}\left(\frac{\sin\theta}{\lambda'/\lambda - \cos\theta}\right) = \tan^{-1}\left(\frac{\sin 60^\circ}{4 - \cos 60^\circ}\right) = 13.9^\circ. \quad (1.194)$$

Problem 1.8

In a double-slit experiment with a source of monoenergetic electrons, detectors are placed along a vertical screen parallel to the y-axis to monitor the diffraction pattern of the electrons emitted from the two slits. When only one slit is open, the amplitude of the electrons detected on the screen is $\psi_1(y,t) = A_1 e^{-i(ky-\omega t)}/\sqrt{1+y^2}$, and when only the other is open the amplitude is $\psi_2(y,t) = A_2 e^{-i(ky+\pi y-\omega t)}/\sqrt{1+y^2}$, where A_1 and A_2 are normalization constants that need to be found. Calculate the intensity detected on the screen when

(a) both slits are open and a light source is used to determine which of the slits the electron went through and

(b) both slits are open and no light source is used.

Plot the intensity registered on the screen as a function of y for cases (a) and (b).

Solution

Using the integral $\int_{-\infty}^{+\infty} dy/(1+y^2) = \pi$, we can obtain the normalization constants at once: $A_1 = A_2 = 1/\sqrt{\pi}$; hence ψ_1 and ψ_2 become $\psi_1(y,t) = e^{-i(ky-\omega t)}/\sqrt{\pi(1+y^2)}$, $\psi_2(y,t) = e^{-i(ky+\pi y-\omega t)}/\sqrt{\pi(1+y^2)}$.

(a) When we use a light source to observe the electrons as they exit from the two slits on their way to the vertical screen, the total intensity recorded on the screen will be determined by a simple addition of the probability densities (or of the separate intensities):

$$I(y) = |\psi_1(y,t)|^2 + |\psi_2(y,t)|^2 = \frac{2}{\pi(1+y^2)}. \quad (1.195)$$

As depicted in Figure 1.19a, the shape of the total intensity displays no interference pattern. Intruding on the electrons with the light source, we distort their motion.

(b) When no light source is used to observe the electrons, the motion will not be distorted and the total intensity will be determined by an addition of the amplitudes, not the intensities:

$$\begin{aligned} I(y) &= |\psi_1(y,t) + \psi_2(y,t)|^2 = \frac{1}{\pi(1+y^2)}\left|e^{-i(ky-\omega t)} + e^{-i(ky+\pi y-\omega t)}\right|^2 \\ &= \frac{1}{\pi(1+y^2)}\left(1+e^{i\pi y}\right)\left(1+e^{-i\pi y}\right) \\ &= \frac{4}{\pi(1+y^2)}\cos^2\left(\frac{\pi}{2}y\right). \end{aligned} \quad (1.196)$$

The shape of this intensity does display an interference pattern which, as shown in Figure 1.19b, results from an oscillating function, $\cos^2(\pi y/2)$, modulated by $4/[\pi(1+y^2)]$.

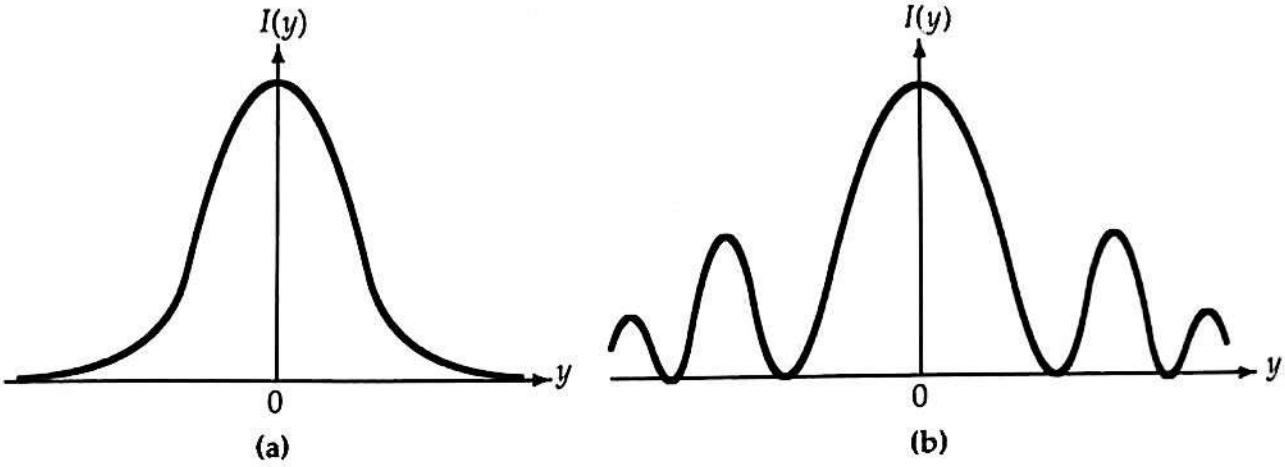


Figure 1.19: Shape of the total intensity generated in a double slit experiment when both slits are open and **(a) a light source is used** to observe the electrons' motion, $I(y) = 2/\pi(1 + y^2)$, and no interference is registered; **(b) no light source is used,** $I(y) = 4/[\pi(1 + y^2)]\cos^2(\pi y/2)$, and an interference pattern occurs.

Problem 1.9
Consider a head-on collision between an α-particle and a lead nucleus. Neglecting the recoil of the lead nucleus, calculate the distance of closest approach of a 9.0 MeV α-particle to the nucleus.

Solution
In this head-on collision the distance of closest approach r_0 can be obtained from the conservation of energy $E_i = E_f$, where E_i is the initial energy of the system, α-particle plus the lead nucleus, when the particle and the nucleus are far from each other and thus feel no electrostatic potential between them. Assuming the lead nucleus to be at rest, E_i is simply the energy of the α-particle: $E_i = 9.0\ \text{MeV} = 9 \times 10^6 \times 1.6 \times 10^{-19}$ J.

As for E_f, it represents the energy of the system when the α-particle is at its closest distance from the nucleus. At this position, the α-particle is at rest and hence has no kinetic energy. The only energy the system has is the electrostatic potential energy between the α-particle and the lead nucleus, which has a positive charge of $82e$. Neglecting the recoil of the lead nucleus and since the charge of the α-particle is positive and equal to $2e$, we have $E_f = (2e)(82e)/(4\pi\varepsilon_0 r_0)$. The energy conservation $E_i = E_f$ or $(2e)(82e)/(4\pi\varepsilon_0 r_0) = E_i$ leads at once to

$$r_0 = \frac{(2e)(82e)}{4\pi\varepsilon_0 E_i} = 2.62 \times 10^{-14}\ \text{m}, \tag{1.197}$$

where we used the values $e = 1.6 \times 10^{-19}$ C and $1/(4\pi\varepsilon_0) = 8.9 \times 10^9$ N m^2 C^{-2}.

Problem 1.10
Considering that a quintuply ionized carbon ion, C^{5+}, behaves like a hydrogen atom, calculate

(a) the radius r_n and energy E_n for a given state n and compare them with the corresponding expressions for hydrogen,

(b) the ionization energy of C^{5+} when it is in its first excited state and compare it with the corresponding value for hydrogen, and

(c) the wavelength corresponding to the transition from state $n = 3$ to state $n = 1$; compare it with the corresponding value for hydrogen.

Solution

(a) The C^{5+} ion is generated by removing five electrons from the carbon atom. To find the expressions for r_{n_C} and E_{n_C} for the C^{5+} ion (which has 6 protons), we need simply to insert $Z = 6$ into (1.81):

$$r_{n_C} = \frac{a_0}{6}n^2, \qquad E_{n_C} = -\frac{36\mathcal{R}}{n^2}, \tag{1.198}$$

where we have dropped the term m_e/M, since it is too small compared to one. Clearly, these expressions are related to their hydrogen counterparts by

$$r_{n_C} = \frac{a_0}{6}n^2 = \frac{r_{n_H}}{6}, \qquad E_{n_C} = -\frac{36\mathcal{R}}{n^2} = 36E_{n_H}. \tag{1.199}$$

(b) The ionization energy is the one needed to remove the only remaining electron of the C^{5+} ion. When the C^{5+} ion is in its first excited state, the ionization energy is

$$E_{2_C} = -\frac{36\mathcal{R}}{4} = -9 \times 13.6\ \text{eV} = -122.4\ \text{eV}, \tag{1.200}$$

which is equal to 36 times the energy needed to ionize the hydrogen atom in its first excited state: $E_{2_H} = -3.4$ eV (note that we have taken $n = 2$ to correspond to the first excited state; as a result, the cases $n = 1$ and $n = 3$ will correspond to the ground and second excited states, respectively).

(c) The wavelength corresponding to the transition from state $n = 3$ to state $n = 1$ can be inferred from the relation $hc/\lambda = E_{3_C} - E_{1_C}$ which, when combined with $E_{1_C} = -489.6$ eV and $E_{3_C} = -54.4$ eV, leads to

$$\lambda = \frac{hc}{E_{3_C} - E_{1_C}} = \frac{2\pi\hbar c}{E_{3_C} - E_{1_C}} = \frac{2\pi 197.33 \times 10^{-9}\ \text{eV m}}{-54.4\ \text{eV} + 489.6\ \text{eV}} = 2.85\ \text{nm}. \tag{1.201}$$

Problem 1.11

(a) Find the Fourier transform for $\phi(k) = \begin{cases} A(a - |k|), & |k| \le a, \\ 0, & |k| > a. \end{cases}$

where a is a positive parameter and A is a normalization factor to be found.

(b) Calculate the uncertainties Δx and Δp and check whether they satisfy the uncertainty principle.

Solution

(a) The normalization factor A can be found at once:

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} |\phi(k)|^2 dk = |A|^2 \int_{-a}^{0} (a + k)^2 dk + |A|^2 \int_{0}^{a} (a - k)^2 dk \\ &= 2|A|^2 \int_0^a (a - k)^2 dk = 2|A|^2 \int_0^a \left(a^2 - 2ak + k^2\right) dk \\ &= \frac{2a^3}{3}|A|^2, \end{aligned} \tag{1.202}$$

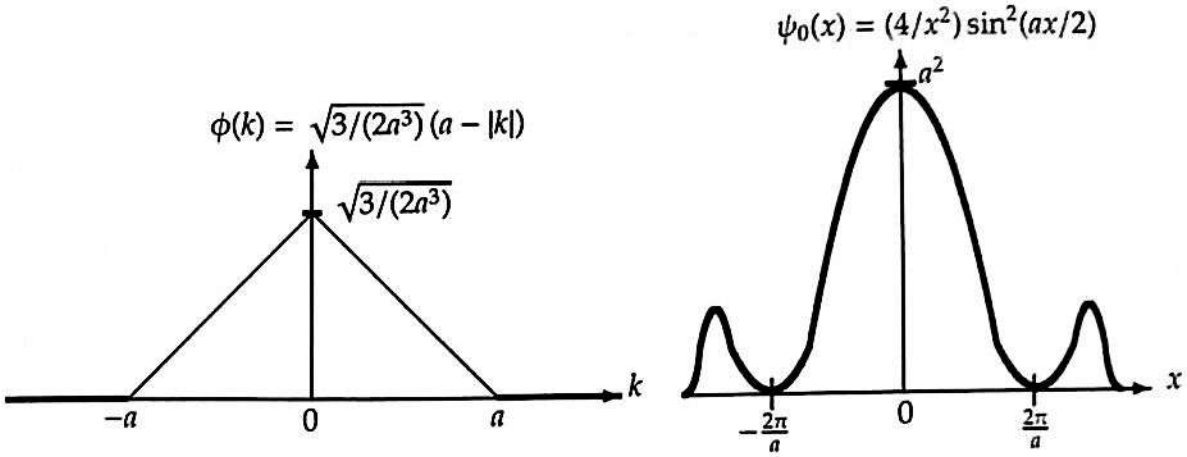


Figure 1.20: The shape of the function $\phi(k)$ and its Fourier transform $\psi_0(x)$.

which yields $A = \sqrt{3/(2a^3)}$. The shape of $\phi(k) = \sqrt{3/(2a^3)}\,(a - |k|)$ is displayed in Figure 1.20.

Now, the Fourier transform of $\phi(k)$ is

$$\begin{aligned}
\psi_0(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{ikx} dk \\
&= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{3}{2a^3}} \left[\int_{-a}^{0} (a+k)\, e^{ikx} dk + \int_{0}^{a} (a-k)\, e^{ikx} dk \right] \\
&= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{3}{2a^3}} \left[\int_{-a}^{0} k e^{ikx} dk - \int_{0}^{a} k e^{ikx} dk + a \int_{-a}^{a} e^{ikx} dk \right].
\end{aligned} \tag{1.203}$$

Using the integrations

$$\int_{-a}^{0} k e^{ikx} dk = \frac{a}{ix} e^{-iax} + \frac{1}{x^2}\left(1 - e^{-iax}\right), \tag{1.204}$$

$$\int_{0}^{a} k e^{ikx} dk = \frac{a}{ix} e^{iax} + \frac{1}{x^2}\left(e^{iax} - 1\right), \tag{1.205}$$

$$\int_{-a}^{a} e^{ikx} dk = \frac{1}{ix}\left(e^{iax} - e^{-iax}\right) = \frac{2\sin(ax)}{x}, \tag{1.206}$$

and after some straightforward calculations, we end up with

$$\psi_0(x) = \frac{4}{x^2} \sin^2\left(\frac{ax}{2}\right). \tag{1.207}$$

As shown in Figure 1.20, this wave packet is localized: it peaks at $x = 0$ and decreases gradually as x increases. We can verify that the maximum of $\psi_0(x)$ occurs at $x = 0$; writing $\psi_0(x)$ as $a^2(ax/2)^{-2}\sin^2(ax/2)$ and since $\lim_{x\to 0} \sin(bx)/(bx) \to 1$, we obtain $\psi_0(0) = a^2$.

(b) Figure 1.20a is quite suggestive in defining the half-width of $\phi(k)$: $\Delta k = a$ (hence the momentum uncertainty is $\Delta p = \hbar a$). By defining the width as $\Delta k = a$, we know with full certainty that the particle is located between $-a \le k \le a$; according to Figure 1.20a, the probability of finding the particle outside this interval is zero, for $\phi(k)$ vanishes when $|k| > a$.

Now, let us find the width Δx of $\psi_0(x)$. Since $\sin(a\pi/2a) = 1$, $\psi_0(\pi/a) = 4a^2/\pi^2$, and that $\psi_0(0) = a^2$, we can obtain from (1.207) that $\psi_0(\pi/a) = 4a^2/\pi^2 = 4/\pi^2\psi_0(0)$, or

$$\frac{\psi_0(\pi/a)}{\psi_0(0)} = \frac{4}{\pi^2}. \tag{1.208}$$

This suggests that $\Delta x = \pi/a$: when $x = \pm\Delta x = \pm\pi/a$ the wave packet $\psi_0(x)$ drops to $4/\pi^2$ from its maximum value $\psi_0(0) = a^2$. In sum, we have $\Delta x = \pi/a$ and $\Delta k = a$; hence

$$\Delta x \Delta k = \pi \tag{1.209}$$

or

$$\Delta x \Delta p = \pi\hbar, \tag{1.210}$$

since $\Delta k = \Delta p/\hbar$. In addition to satisfying Heisenberg's uncertainty principle (1.62), this relation shows that the product $\Delta x\Delta p$ is higher than $\hbar/2$: $\Delta x\Delta p > \hbar/2$. The wave packet (1.207) therefore offers a clear illustration of the general statement outlined above; namely, only Gaussian wave packets yield the *lowest* limit to Heisenberg's uncertainty principle $\Delta x\Delta p = \hbar/2$ (see (1.119)). All other wave packets, such as (1.207), yield higher values for the product $\Delta x\Delta p$.

Problem 1.12
Calculate the group and phase velocities for the wave packet corresponding to a relativistic particle.

Solution
Recall that the energy and momentum of a relativistic particle are given by

$$E = mc^2 = \frac{m_0c^2}{\sqrt{1 - v^2/c^2}}, \qquad p = mv = \frac{m_0v}{\sqrt{1 - v^2/c^2}}, \tag{1.211}$$

where m_0 is the rest mass of the particle and c is the speed of light in a vacuum. Squaring and adding the expressions of E and p, we obtain $E^2 = p^2c^2 + m_0^2c^4$; hence

$$E = c\sqrt{p^2 + m_0^2c^2}. \tag{1.212}$$

Using this relation along with $p^2 + m_0^2c^2 = m_0^2c^2/(1 - v^2/c^2)$ and (1.127), we can show that the group velocity is given as follows:

$$v_g = \frac{dE}{dp} = \frac{d}{dp}\left(c\sqrt{p^2 + m_0^2c^2}\right) = \frac{pc}{\sqrt{p^2 + m_0^2c^2}} = v. \tag{1.213}$$

The group velocity is thus equal to the speed of the particle, $v_g = v$.

The phase velocity can be found from (1.127) and (1.212): $v_{ph} = E/p = c\sqrt{1 + m_0^2c^2/p^2}$ which, when combined with $p = m_0v/\sqrt{1 - v^2/c^2}$, leads to $\sqrt{1 + m_0^2c^2/p^2} = c/v$; hence

$$v_{ph} = \frac{E}{p} = c\sqrt{1 + \frac{m_0^2c^2}{p^2}} = \frac{c^2}{v}. \tag{1.214}$$

This shows that the phase velocity of the wave corresponding to a relativistic particle with $m_0 \neq 0$ is larger than the speed of light, $v_{ph} = c^2/v > c$. This is indeed unphysical. The result $v_{ph} > c$ seems to violate the special theory of relativity, which states that the speed of material particles cannot exceed c. In fact, this principle is not violated because v_{ph} does not represent the velocity of the particle; the velocity of the particle is represented by the group velocity (1.213). As a result, the phase speed of a relativistic particle has no meaningful physical significance.

Finally, the product of the group and phase velocities is equal to c^2, i.e., $v_g v_{ph} = c^2$.

Problem 1.13

The angular frequency of the surface waves in a liquid is given in terms of the wave number k by $\omega = \sqrt{gk + Tk^3/\rho}$, where g is the acceleration due to gravity, ρ is the density of the liquid, and T is the surface tension (which gives an upward force on an element of the surface liquid). Find the phase and group velocities for the limiting cases when the surface waves have: (a) very large wavelengths and (b) very small wavelengths.

Solution

The phase velocity can be found at once from (1.124):

$$v_{ph} = \frac{\omega}{k} = \sqrt{\frac{g}{k} + \frac{T}{\rho}k} = \sqrt{\frac{g\lambda}{2\pi} + \frac{2\pi T}{\rho\lambda}}, \tag{1.215}$$

where we have used the fact that $k = 2\pi/\lambda$, λ being the wavelength of the surface waves.

(a) If λ is very large, we can neglect the second term in (1.215); hence

$$v_{ph} = \sqrt{\frac{g\lambda}{2\pi}} = \sqrt{\frac{g}{k}}. \tag{1.216}$$

In this approximation the phase velocity does not depend on the nature of the liquid, since it depends on no parameter pertaining to the liquid such as its density or surface tension. This case corresponds, for instance, to deepwater waves, called gravity waves.

To obtain the group velocity, let us differentiate (1.216) with respect to k: $dv_{ph}/dk = -(1/2k)\sqrt{g/k} = -v_{ph}/2k$. A substitution of this relation into (1.125) shows that the group velocity is half the phase velocity:

$$v_g = \frac{d\omega}{dk} = v_{ph} + k\frac{dv_{ph}}{dk} = v_{ph} - \frac{1}{2}v_{ph} = \frac{1}{2}v_{ph} = \frac{1}{2}\sqrt{\frac{g\lambda}{2\pi}}. \tag{1.217}$$

The longer the wavelength, the faster the group velocity. This explains why a strong, steady wind will produce waves of longer wavelength than those produced by a swift wind.

(b) If λ is very small, the second term in (1.215) becomes the dominant one. So, retaining only the second term, we have

$$v_{ph} = \sqrt{\frac{2\pi T}{\rho\lambda}} = \sqrt{\frac{T}{\rho}k}, \tag{1.218}$$

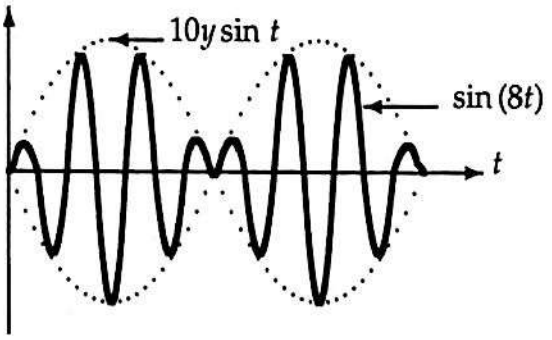


Figure 1.21: Shape of the wave packet $\psi(y,t) = 10y \sin t \sin 8t$. The function $\sin 8t$, the solid curve, is modulated by $10y \sin t$, the dashed curve.

which leads to $dv_{ph}/dk = \sqrt{Tk/\rho}/2k = v_{ph}/2k$. Inserting this expression into (1.125), we obtain the group velocity

$$v_g = v_{ph} + k\frac{dv_{ph}}{dk} = v_{ph} + \frac{1}{2}v_{ph} = \frac{3}{2}v_{ph}; \tag{1.219}$$

hence the smaller the wavelength, the faster the group velocity. These are called ripple waves; they occur, for instance, when a container is subject to vibrations of high frequency and small amplitude or when a gentle wind blows on the surface of a fluid.

Problem 1.14

This problem is designed to illustrate the superposition principle and the concepts of modulated and modulating functions in a wave packet. Consider two wave functions $\psi_1(y,t) = 5y\cos 7t$ and $\psi_2(y,t) = -5y\cos 9t$, where y and t are in meters and seconds, respectively. Show that their superposition generates a wave packet. Plot it and identify the modulated and modulating functions.

Solution

Using the relation $\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$, we can write the superposition of $\psi_1(y,t)$ and $\psi_2(y,t)$ as follows:

$$\begin{aligned} \psi(y,t) &= \psi_1(y,t) + \psi_2(y,t) = 5y\cos 7t - 5y\cos 9t \\ &= 5y(\cos 8t\cos t + \sin 8t\sin t) - 5y(\cos 8t\cos t - \sin 8t\sin t) \\ &= 10y\sin t\sin 8t. \end{aligned} \tag{1.220}$$

The periods of $10y\sin t$ and $\sin(8t)$ are given by 2π and $2\pi/8$, respectively. Since the period of $10y\sin t$ is larger than that of $\sin 8t$, $10y\sin t$ must be the modulating function and $\sin 8t$ the modulated function. As depicted in Figure 1.21, we see that $\sin 8t$ is modulated by $10y\sin t$.

Problem 1.15

(a) Calculate the final size of the wave packet representing a free particle after traveling a distance of 100 m for the following four cases where the particle is

(i) a 25 eV electron whose wave packet has an initial width of 10^{-6} m,

(ii) a 25 eV electron whose wave packet has an initial width of 10^{-8} m,

(iii) a 100 MeV electron whose wave packet has an initial width of 1 mm, and

(iv) a 100 g object of size 1 cm moving at a speed of 50 m s^{-1}.

(b) Estimate the times required for the wave packets of the electron in (i) and the object in (iv) to spread to 10 mm and 10 cm, respectively. Discuss the results obtained.

Solution

(a) If the initial width of the wave packet of the particle is Δx_0, the width at time t is given by

$$\Delta x(t) = \Delta x_0 \sqrt{1 + \left(\frac{\delta x}{\Delta x_0}\right)^2}, \tag{1.221}$$

where the dispersion factor is given by

$$\frac{\delta x}{\Delta x_0} = \frac{2\hbar t}{ma^2} = \frac{\hbar t}{2m(a/2)^2} = \frac{\hbar t}{2m\,(\Delta x_0)^2}. \tag{1.222}$$

(i) For the 25 eV electron, which is clearly not relativistic, the time to travel the $L = 100$ m distance is given by $t = L/v = L\sqrt{mc^2/2E}/c$, since $E = \frac{1}{2}mv^2 = \frac{1}{2}mc^2(v^2/c^2)$ or $v = c\sqrt{2E/(mc^2)}$. We can therefore write the dispersion factor as

$$\frac{\delta x}{\Delta x_0} = \frac{\hbar}{2m\Delta x_0^2}t = \frac{\hbar}{2m\Delta x_0^2}\frac{L}{c}\sqrt{\frac{mc^2}{2E}} = \frac{\hbar cL}{2mc^2\Delta x_0^2}\sqrt{\frac{mc^2}{2E}}. \tag{1.223}$$

The numerics of this expression can be made easy by using the following quantities: $\hbar c \simeq 197 \times 10^{-15}$ MeV m, the rest mass energy of an electron is $mc^2 = 0.5$ MeV, $\Delta x_0 = 10^{-6}$ m, $E = 25$ eV $= 25 \times 10^{-6}$ MeV, and $L = 100$ m. Inserting these quantities into (1.223), we obtain

$$\frac{\delta x}{\Delta x_0} \simeq \frac{197 \times 10^{-15}\ \text{MeV m} \times 100\ \text{m}}{2 \times 0.5\ \text{MeV} \times 10^{-12}\ \text{m}^2}\sqrt{\frac{0.5\ \text{MeV}}{2 \times 25 \times 10^{-6}\ \text{MeV}}} \simeq 2 \times 10^3; \tag{1.224}$$

the time it takes the electron to travel the 100 m distance is given, as shown above, by

$$t = \frac{L}{c}\sqrt{\frac{mc^2}{2E}} = \frac{100\ \text{m}}{3 \times 10^8\ \text{m s}^{-1}}\sqrt{\frac{0.5\ \text{MeV}}{2 \times 25 \times 10^{-6}\ \text{MeV}}} = 3.3 \times 10^{-5}\ \text{s}. \tag{1.225}$$

Using $t = 3.3 \times 10^{-5}$ s and substituting (1.224) into (1.221), we obtain

$$\Delta x(t = 3.3 \times 10^{-5}\ \text{s}) = 10^{-6}\ \text{m} \times \sqrt{1 + 4 \times 10^6} \simeq 2 \times 10^{-3}\ \text{m} = 2\ \text{mm}. \tag{1.226}$$

The width of the wave packet representing the electron has increased from an initial value of 10^{-6} m to 2×10^{-3} m, i.e., by a factor of about 10^3. The spread of the electron's wave packet is thus quite large.

(ii) The calculation needed here is identical to that of part (i), except the value of Δx_0 is now 10^{-8} m instead of 10^{-6} m. This leads to $\delta x/\Delta x_0 \simeq 2 \times 10^7$ and hence the width is $\Delta x(t) = 20$ cm; the width has therefore increased by a factor of about 10^7. This calculation is intended to show that *the narrower the initial wave packet, the larger the final spread.* In fact, starting in part (i) with an initial width of 10^{-6} m, the final width has increased to 2×10^{-3} m by a factor of about 10^3; but in part (ii) we started with an initial width of 10^{-8} m, and the final width has increased to 20 cm by a factor of about 10^7.

(iii) The motion of a 100 MeV electron is relativistic; hence to good approximation, its speed is equal to the speed of light, $v \simeq c$. Therefore the time it takes the electron to travel a distance of $L = 100$ m is $t \simeq L/c = 3.3 \times 10^{-7}$ s. The dispersion factor for this electron can be obtained from (1.222) where $\Delta x_0 = 10^{-3}$ m:

$$\frac{\delta x}{\Delta x_0} = \frac{\hbar L}{2mc\Delta x_0^2} = \frac{\hbar c L}{2mc^2\Delta x_0^2} \simeq \frac{197 \times 10^{-15}\,\text{MeV m} \times 100\,\text{m}}{2 \times 0.5\,\text{MeV} \times 10^{-6}\,\text{m}^2} \simeq 2 \times 10^{-5}. \tag{1.227}$$

The increase in the width of the wave packet is relatively small:

$$\Delta x(t = 3.3 \times 10^{-7}\,\text{s}) = 10^{-3}\,\text{m} \times \sqrt{1 + 4 \times 10^{-10}} \simeq 10^{-3}\,\text{m} = \Delta x_0. \tag{1.228}$$

So the width did not increase appreciably. We can conclude from this calculation that, when the motion of a microscopic particle is relativistic, the width of the corresponding wave packet increases by a relatively small amount.

(iv) In the case of a macroscopic object of mass $m = 0.1$ kg, the time to travel the distance $L = 100$ m is $t = L/v = 100\,\text{m}/50\,\text{m s}^{-1} = 2$ s. Since the size of the system is about $\Delta x_0 = 1$ cm $= 0.01$ m and $\hbar = 1.05 \times 10^{-34}$ J s, the dispersion factor for the object can be obtained from (1.222):

$$\frac{\delta x}{\Delta x_0} = \frac{\hbar t}{2m\Delta x_0^2} \simeq \frac{1.05 \times 10^{-34}\,\text{J s} \times 2\,\text{s}}{2 \times 0.1\,\text{kg} \times 10^{-4}\,\text{m}^2} \simeq 10^{-29}. \tag{1.229}$$

Since $\delta x/\Delta x_0 = 10^{-29} \ll 1$, the increase in the width of the wave packet is utterly undetectable:

$$\Delta x(2s) = 10^{-2}\,\text{m} \times \sqrt{1 + 10^{-58}} \simeq 10^{-2}\,\text{m} = \Delta x_0. \tag{1.230}$$

(b) Using (1.221) and (1.222) we obtain the expression for the time t in which the wave packet spreads to $\Delta x(t)$:

$$t = \tau \sqrt{\left(\frac{\Delta x(t)}{\Delta x_0}\right)^2 - 1}, \tag{1.231}$$

where τ represents a time constant $\tau = 2m(\Delta x_0)^2/\hbar$ (see (1.153)). The time constant for the electron of part (i) is given by

$$\tau = \frac{2mc^2(\Delta x_0)^2}{\hbar c^2} \simeq \frac{2 \times 0.5\,\text{MeV} \times 10^{-12}\,\text{m}^2}{197 \times 10^{-15}\,\text{MeV m} \times 3 \times 10^8\,\text{m s}^{-1}} = 1.7 \times 10^{-8}\,\text{s}, \tag{1.232}$$

and the time constant for the object of part (iv) is given by

$$\tau = \frac{2m(\Delta x_0)^2}{\hbar} \simeq \frac{2 \times 0.1\,\text{kg} \times 10^{-4}\,\text{m}^2}{1.05 \times 10^{-34}\,\text{J s}} = 1.9 \times 10^{29}\,\text{s}. \tag{1.233}$$

Note that the time constant, while very small for a microscopic particle, is exceedingly large for macroscopic objects.

On the one hand, a substitution of the time constant (1.232) into (1.231) yields the time required for the electron's packet to spread to 10 mm:

$$t = 1.7 \times 10^{-8}\,\text{s} \sqrt{\left(\frac{10^{-2}}{10^{-6}}\right)^2 - 1} \simeq 1.7 \times 10^{-4}\,\text{s}. \tag{1.234}$$

On the other hand, a substitution of (1.233) into (1.231) gives the time required for the object to spread to 10 cm:

$$t = 1.9 \times 10^{29}\ \text{s} \sqrt{\left(\frac{10^{-1}}{10^{-2}}\right)^2 - 1} \simeq 1.9 \times 10^{30}\ \text{s}. \tag{1.235}$$

The result (1.234) shows that the size of the electron's wave packet grows in a matter of 1.7×10^{-4} s from 10^{-6} m to 10^{-2} m, a very large spread in a very short time. As for (1.235), it shows that the object has to be constantly in motion for about 1.9×10^{30} s for its wave packet to grow from 1 cm to 10 cm, a small spread for such an absurdly large time; this time is absurd because it is much larger than the age of the Universe, which is about 4.7×10^{17} s. We see that the spread of macroscopic objects becomes appreciable only if the motion lasts for a long, long time. However, the spread of microscopic objects is fast and large.

We can summarize these ideas in three points:

- The width of the wave packet of a nonrelativistic, microscopic particle increases substantially and quickly. The narrower the wave packet at the start, the further and the quicker it will spread.
- When the particle is microscopic and relativistic, the width corresponding to its wave packet does not increase appreciably.
- For a nonrelativistic, macroscopic particle, the width of its corresponding wave packet remains practically constant. The spread becomes appreciable only after absurdly long times, times that are larger than the lifetime of the Universe itself!

Problem 1.16

A neutron is confined in space to 10^{-14} m. Calculate the time its packet will take to spread to
(a) four times its original size,
(b) a size equal to the Earth's diameter, and
(c) a size equal to the distance between the Earth and the Moon.

Solution

Since the rest mass energy of a neutron is equal to $m_n c^2 = 939.6$ MeV, we can infer the time constant for the neutron from (1.232):

$$\tau = \frac{2m_n c^2 (\Delta x_0)^2}{\hbar c^2} \simeq \frac{2 \times 939.6\ \text{MeV} \times (10^{-14}\ \text{m})^2}{197 \times 10^{-15}\ \text{MeV m} \times 3 \times 10^8\ \text{m s}^{-1}} = 3.2 \times 10^{-21}\ \text{s}. \tag{1.236}$$

Inserting this value in (1.231) we obtain the time it takes for the neutron's packet to grow from an initial width Δx_0 to a final size $\Delta x(t)$:

$$t = \tau \sqrt{\left(\frac{\Delta x(t)}{\Delta x_0}\right)^2 - 1} = 3.2 \times 10^{-21}\ \text{s} \sqrt{\left(\frac{\Delta x(t)}{\Delta x_0}\right)^2 - 1}. \tag{1.237}$$

The calculation of t reduces to simple substitutions.

(a) Substituting $\Delta x(t) = 4\Delta x_0$ into (1.237), we obtain the time needed for the neutron's packet to expand to four times its original size:

$$t = 3.2 \times 10^{-21}\ \text{s} \sqrt{16 - 1} = 1.2 \times 10^{-20}\ \text{s}. \tag{1.238}$$

(b) The neutron's packet will expand from an initial size of 10^{-14} m to 12.7×10^6 m (the diameter of the Earth) in a time of

$$t = 3.2 \times 10^{-21}\ \text{s} \sqrt{\left(\frac{12.7 \times 10^6\ \text{m}}{10^{-14}\ \text{m}}\right)^2 - 1} = 4.1\ \text{s}. \tag{1.239}$$

(c) The time needed for the neutron's packet to spread from 10^{-14} m to 3.84×10^8 m (the distance between the Earth and the Moon) is

$$t = 3.2 \times 10^{-21}\ \text{s} \sqrt{\left(\frac{3.84 \times 10^8\ \text{m}}{10^{-14}\ \text{m}}\right)^2 - 1} = 12.3\ \text{s}. \tag{1.240}$$

The calculations carried out in this problem show that the spread of the packets of microscopic particles is significant and occurs very fast: the size of the packet for an earthly neutron can expand to reach the Moon in a mere 12.3 s! Such an immense expansion in such a short time is indeed hard to visualize classically. One should not confuse the packet's expansion with a growth in the size of the system. As mentioned above, the spread of the wave packet does not mean that the material particle becomes bloated. It simply implies a redistribution of the probability density. In spite of the significant spread of the wave packet, the packet's norm is always conserved; as shown in (1.154) it is equal to 1.

Problem 1.17
Use the uncertainty principle to estimate: (a) the ground state radius of the hydrogen atom and (b) the ground state energy of the hydrogen atom.

Solution
(a) According to the uncertainty principle, the electron's momentum and the radius of its orbit are related by $rp \sim \hbar$; hence $p \sim \hbar/r$. To find the ground state radius, we simply need to minimize the electron–proton energy

$$E(r) = \frac{p^2}{2m_e} - \frac{e^2}{4\pi\varepsilon_0 r} = \frac{\hbar^2}{2m_e r^2} - \frac{e^2}{4\pi\varepsilon_0 r} \tag{1.241}$$

with respect to r:

$$0 = \frac{dE}{dr} = -\frac{\hbar^2}{m_e r_0^3} + \frac{e^2}{4\pi\varepsilon_0 r_0^2}. \tag{1.242}$$

This leads to the Bohr radius

$$r_0 = \frac{4\pi\varepsilon_0\hbar^2}{m_e e^2} = 0.053\ \text{nm}. \tag{1.243}$$

(b) Inserting (1.243) into (1.241), we obtain the Bohr energy:

$$E(r_0) = \frac{\hbar^2}{2mr_0^2} - \frac{e^2}{4\pi\varepsilon_0 r_0} = -\frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 = -13.6\ \text{eV}. \tag{1.244}$$

The results obtained for r_0 and $E(r_0)$, as shown in (1.243) and (1.244), are indeed impressively accurate given the crudeness of the approximation.

Problem 1.18

Consider the bound state of two quarks having the same mass m and interacting via a potential energy $V(r) = kr$ where k is a constant.

(a) Using the Bohr model, find the speed, the radius, and the energy of the system in the case of circular orbits. Determine also the angular frequency of the radiation generated by a transition of the system from energy state n to energy state m.

(b) Obtain numerical values for the speed, the radius, and the energy for the case of the ground state, $n = 1$, by taking a quark mass of $mc^2 = 2$ GeV and $k = 0.5$ GeV fm^{-1}.

Solution

(a) Consider the two quarks to move circularly, much like the electron and proton in a hydrogen atom; then we can write the force between them as

$$\mu \frac{v^2}{r} = \frac{dV(r)}{dr} = k, \tag{1.245}$$

where $\mu = m/2$ is the reduced mass and $V(r)$ is the potential. From the Bohr quantization condition of the orbital angular momentum, we have

$$L = \mu v r = n\hbar. \tag{1.246}$$

Multiplying (1.245) by (1.246), we end up with $\mu^2 v^3 = n\hbar k$, which yields the (quantized) speed of the relative motion for the two-quark system:

$$v_n = \left(\frac{\hbar k}{\mu^2}\right)^{1/3} n^{1/3}. \tag{1.247}$$

The radius can be obtained from (1.246), $r_n = n\hbar/(\mu v_n)$; using (1.247), this leads to

$$r_n = \left(\frac{\hbar^2}{\mu k}\right)^{1/3} n^{2/3}. \tag{1.248}$$

We can obtain the total energy of the relative motion by adding the kinetic and potential energies:

$$E_n = \frac{1}{2}\mu v_n^2 + k r_n = \frac{3}{2}\left(\frac{\hbar^2 k^2}{\mu}\right)^{1/3} n^{2/3}. \tag{1.249}$$

In deriving this relation, we have used the relations for v_n and r_n as given by (1.247) by (1.248), respectively.

The angular frequency of the radiation generated by a transition from n to m is given by

$$\omega_{nm} = \frac{E_n - E_m}{\hbar} = \frac{3}{2}\left(\frac{k^2}{\mu\hbar}\right)^{1/3}\left(n^{2/3} - m^{2/3}\right). \tag{1.250}$$

(b) Inserting $n = 1$, $\hbar c \simeq 0.197$ GeV fm, $\mu c^2 = mc^2/2 = 1$ GeV, and $k = 0.5$ GeV fm^{-1} into (1.247) to (1.249), we have

$$v_1 = \left(\frac{\hbar c k}{(\mu c^2)^2}\right)^{1/3} c \simeq \left(\frac{0.197 \text{ GeV fm} \times 0.5 \text{ GeV fm}^{-1}}{(1 \text{ GeV})^2}\right)^{1/3} c = 0.46c, \tag{1.251}$$

where c is the speed of light and

$$r_1 = \left(\frac{(\hbar c)^2}{\mu c^2 k}\right)^{1/3} \simeq \left(\frac{(0.197\ \text{GeV fm})^2}{1\ \text{GeV} \times 0.5\ \text{GeV fm}^{-1}}\right)^{1/3} = 0.427\ \text{fm}, \tag{1.252}$$

$$E_1 = \frac{3}{2}\left(\frac{(\hbar c)^2 k^2}{\mu c^2}\right)^{1/3} \simeq \frac{3}{2}\left(\frac{(0.197\ \text{GeV fm})^2 (0.5\ \text{GeV fm}^{-1})^2}{1\ \text{GeV}}\right)^{1/3} = 0.32\ \text{GeV}. \tag{1.253}$$

1.11 Exercises

Exercise 1.1
Consider a metal that is being welded.
(a) How hot is the metal when it radiates most strongly at 490 nm?
(b) Assuming that it radiates like a blackbody, calculate the intensity of its radiation.

Exercise 1.2
Consider a star, a light bulb, and a slab of ice; their respective temperatures are 8500 K, 850 K, and 273.15 K.
(a) Estimate the wavelength at which their radiated energies peak.
(b) Estimate the intensities of their radiation.

Exercise 1.3
Consider a 75 W light bulb and an 850 W microwave oven. If the wavelengths of the radiation they emit are 500 nm and 150 mm, respectively, estimate the number of photons they emit per second. Are the quantum effects important in them?

Exercise 1.4
Assuming that a given star radiates like a blackbody, estimate
(a) the temperature at its surface and
(b) the wavelength of its strongest radiation,
when it emits a total intensity of 575 MW m^{-2}.

Exercise 1.5
The intensity reaching the surface of the Earth from the Sun is about 1.36 kW m^{-2}. Assuming the Sun to be a sphere (of radius 6.96×10^8 m) that radiates like a blackbody, estimate
(a) the temperature at its surface and the wavelength of its strongest radiation, and
(b) the total power radiated by the Sun (the Earth–Sun distance is 1.5×10^{11} m).

Exercise 1.6
(a) Calculate: (i) the energy spacing ΔE between the ground state and the first excited state of the hydrogen atom; (ii) and the ratio $\Delta E/E_1$ between the spacing and the ground state energy.
(b) Consider now a macroscopic system: a simple pendulum which consists of a 5 g mass attached to a 2 m long, massless and inextensible string. Calculate (i) the total energy E_1 of the pendulum when the string makes an angle of 60° with the vertical; (ii) the frequency of the pendulum's small oscillations and the energy ΔE of one quantum; and (iii) the ratio $\Delta E/E_1$.
(c) Examine the sizes of the ratio $\Delta E/E_1$ calculated in parts (a) and (b) and comment on the importance of the quantum effects for the hydrogen atom and the pendulum.

Exercise 1.7
A beam of X-rays from a sulfur source ($\lambda = 53.7\,\text{nm}$) and a γ-ray beam from a Cs^{137} sample ($\lambda = 0.19\,\text{nm}$) impinge on a graphite target. Two detectors are set up at angles 30° and 120° from the direction of the incident beams.

(a) Estimate the wavelength shifts of the X-rays and the γ-rays recorded at both detectors.

(b) Find the kinetic energy of the recoiling electron in each of the four cases.

(c) What percentage of the incident photon energy is lost in the collision in each of the four cases?

Exercise 1.8
It has been suggested that high energy photons might be found in cosmic radiation, as a result of the *inverse Compton effect*, i.e., a photon of visible light *gains* energy by scattering from a high energy *proton*. If the proton has a momentum of 10^{10} eV/c, find the maximum final energy of an initially yellow photon emitted by a sodium atom ($\lambda_0 = 2.1\,\text{nm}$).

Exercise 1.9
Estimate the number of photons emitted per second from a 75 W light bulb; use 575 nm as the average wavelength of the (visible) light emitted. Is the quantum nature of this radiation important?

Exercise 1.10
A 0.7 MeV photon scatters from an electron initially at rest. If the photon scatters at an angle of 35°, calculate

(a) the energy and wavelength of the scattered photon,

(b) the kinetic energy of the recoiling electron, and

(c) the angle at which the electron recoils.

Exercise 1.11
Light of wavelength 350 nm is incident on a metallic surface of work function 1.9 eV.

(a) Calculate the kinetic energy of the ejected electrons.

(b) Calculate the cutoff frequency of the metal.

Exercise 1.12
Find the wavelength of the radiation that can eject electrons from the surface of a zinc sheet with a kinetic energy of 75 eV; the work function of zinc is 3.74 eV. Find also the cutoff wavelength of the metal.

Exercise 1.13
If the stopping potential of a metal when illuminated with a radiation of wavelength 480 nm is 1.2 V, find

(a) the work function of the metal,

(b) the cutoff wavelength of the metal, and

(c) the maximum energy of the ejected electrons.

Exercise 1.14
Find the maximum Compton wave shift corresponding to a collision between a photon and a proton at rest.

Exercise 1.15
If the stopping potential of a metal when illuminated with a radiation of wavelength 150 nm is 7.5 V, calculate the stopping potential of the metal when illuminated by a radiation of wavelength 275 nm.

Exercise 1.16
A light source of frequency 9.5×10^{14} Hz illuminates the surface of a metal of work function 2.8 eV and ejects electrons. Calculate
(a) the stopping potential,
(b) the cutoff frequency, and
(c) the kinetic energy of the ejected electrons.

Exercise 1.17
Consider a metal with a cutoff frequency of 1.2×10^{14} Hz.
(a) Find the work function of the metal.
(b) Find the kinetic energy of the ejected electrons when the metal is illuminated with a radiation of frequency 7×10^{14} Hz.

Exercise 1.18
A light of frequency 7.2×10^{14} Hz is incident on four different metallic surfaces of cesium, aluminum, cobalt, and platinum whose work functions are 2.14 eV, 4.08 eV, 3.9 eV, and 6.35 eV, respectively.
(a) Which among these metals will exhibit the photoelectric effect?
(b) For each one of the metals producing photoelectrons, calculate the maximum kinetic energy for the electrons ejected.

Exercise 1.19
Consider a metal with stopping potentials of 9 V and 4 V when illuminated by two sources of frequencies 17×10^{14} Hz and 8×10^{14} Hz, respectively.
(a) Use these data to find a numerical value for the Planck constant.
(b) Find the work function and the cutoff frequency of the metal.
(c) Find the maximum kinetic energy of the ejected electrons when the metal is illuminated with a radiation of frequency 12×10^{14} Hz.

Exercise 1.20
Using energy and momentum conservation requirements, show that a free electron cannot absorb all the energy of a photon.

Exercise 1.21
Photons of wavelength 5 nm are scattered from electrons that are at rest. If the photons scatter at 60° relative to the incident photons, calculate
(a) the Compton wave shift,
(b) the kinetic energy imparted to the recoiling electrons, and
(c) the angle at which the electrons recoil.

Exercise 1.22
X-rays of wavelength 0.0008 nm collide with electrons initially at rest. If the wavelength of the scattered photons is 0.0017 nm, determine

(a) the kinetic energy of the recoiling electrons,
(b) the angle at which the photons scatter, and
(c) the angle at which the electrons recoil.

Exercise 1.23
Photons of energy 0.7 MeV are scattered from electrons initially at rest. If the energy of the scattered photons is 0.5 MeV, find
(a) the wave shift,
(b) the angle at which the photons scatter,
(c) the angle at which the electrons recoil, and
(d) the kinetic energy of the recoiling electrons.

Exercise 1.24
In a Compton scattering of photons from electrons at rest, if the photons scatter at an angle of 45° and if the wavelength of the scattered photons is 9×10^{-13} m, find
(a) the wavelength and the energy of the incident photons,
(b) the energy of the recoiling electrons and the angle at which they recoil.

Exercise 1.25
When scattering photons from electrons at rest, if the scattered photons are detected at 90° and if their wavelength is double that of the incident photons, find
(a) the wavelength of the incident photons,
(b) the energy of the recoiling electrons and the angle at which they recoil, and
(c) the energies of the incident and scattered photons.

Exercise 1.26
In scattering electrons from a crystal, the first maximum is observed at an angle of 60°. What must be the energy of the electrons that will enable us to probe as deep as 19 nm inside the crystal?

Exercise 1.27
Estimate the resolution of a microscope which uses electrons of energy 175 eV.

Exercise 1.28
What are the longest and shortest wavelengths in the Balmer and Paschen series for hydrogen?

Exercise 1.29
(a) Calculate the ground state energy of the doubly ionized lithium ion, Li^{2+}, obtained when one removes two electrons from the lithium atom.
(b) If the lithium ion Li^{2+} is bombarded with a photon and subsequently absorbs it, calculate the energy and wavelength of the photon needed to excite the Li^{2+} ion into its third excited state.

Exercise 1.30
Consider a tenfold ionized sodium ion, Na^{10+}, which is obtained by removing ten electrons from an Na atom.
(a) Calculate the orbiting speed and orbital angular momentum of the electron (with respect to the ion's origin) when the ion is in its fourth excited state.
(b) Calculate the frequency of the radiation emitted when the ion deexcites from its fourth excited state to the first excited state.

Exercise 1.31
Calculate the wavelength of the radiation needed to excite the triply ionized beryllium atom, Be^{3+}, from the ground state to its third excited state.

Exercise 1.32
According to the classical model of the hydrogen atom, an electron moving in a circular orbit of radius 0.053 nm around a proton fixed at the center is unstable, and the electron should eventually collapse into the proton. Estimate how long it would take for the electron to collapse into the proton.

Hint: Start with the classical expression for radiation from an accelerated charge

$$\frac{dE}{dt} = -\frac{2}{3}\frac{e^2a^2}{4\pi\varepsilon_0 c^3}, \qquad E = \frac{p^2}{2m} - \frac{e^2}{4\pi\varepsilon_0 r} = -\frac{e^2}{8\pi\varepsilon_0 r},$$

where a is the *acceleration* of the electron and E is its total energy.

Exercise 1.33
Calculate the de Broglie wavelength of
(a) an electron of kinetic energy 54 eV,
(b) a proton of kinetic energy 70 MeV,
(c) a 100 g bullet moving at 1200 m s^{-1}, and
Useful data: $m_e c^2 = 0.511$ MeV, $m_p c^2 = 938.3$ MeV, $\hbar c \simeq 197.3$ eV nm.

Exercise 1.34
A simple one-dimensional harmonic oscillator is a particle acted upon by a linear restoring force $F(x) = -m\omega^2 x$. Classically, the minimum energy of the oscillator is zero, because we can place it precisely at $x = 0$, its equilibrium position, while giving it zero initial velocity. Quantum mechanically, the uncertainty principle does not allow us to localize the particle precisely and simultaneously have it at rest. Using the uncertainty principle, estimate the minimum energy of the quantum mechanical oscillator.

Exercise 1.35
Consider a double-slit experiment where the waves emitted from the slits superpose on a vertical screen parallel to the y-axis. When only one slit is open, the amplitude of the wave which gets through is $\psi_1(y,t) = e^{-y^2/32}e^{i(\omega t - ay)}$ and when only the other slit is open, the amplitude is $\psi_2(y,t) = e^{-y^2/32}e^{i(\omega t - ay - \pi y)}$.

(a) What is the interference pattern along the y-axis with both slits open? Plot the intensity of the wave as a function of y.

(b) What would be the intensity if we put a light source behind the screen to measure which of the slits the light went through? Plot the intensity of the wave as a function of y.

Exercise 1.36
Consider the following three wave functions:

$$\psi_1(y) = A_1 e^{-y^2}, \qquad \psi_2(y) = A_2 e^{-y^2/2}, \qquad \psi_3(y) = A_3(e^{-y^2} + ye^{-y^2/2}),$$

where A_1, A_2, and A_3 are normalization constants.
(a) Find the constants A_1, A_2, and A_3 so that ψ_1, ψ_2, and ψ_3 are normalized.
(b) Find the probability that each one of the states will be in the interval $-1 < y < 1$.

Exercise 1.37
Find the Fourier transform $\phi(p)$ of the following function and plot it:

$$\psi(x) = \begin{cases} 1 - |x|, & |x| < 1, \\ 0, & |x| \geq 1. \end{cases}$$

Exercise 1.38
(a) Find the Fourier transform of $\phi(k) = Ae^{-a|k|-ibk}$, where a and b are real numbers, but a is positive.
(b) Find A so that $\psi(x)$ is normalized.
(c) Find the x and k uncertainties and calculate the uncertainty product $\Delta x \Delta p$. Does it satisfy Heisenberg's uncertainty principle?

Exercise 1.39
(a) Find the Fourier transform $\psi(x)$ of

$$\phi(p) = \begin{cases} 0, & p < -p_0, \\ A, & -p_0 < p < p_0, \\ 0, & p_0 < p, \end{cases}$$

where A is a real constant.
(b) Find A so that $\psi(x)$ is normalized and plot $\phi(p)$ and $\psi(x)$. *Hint:* The following integral might be needed: $\int_{-\infty}^{+\infty} dx\, (\sin^2(ax))/x^2 = \pi a$.
(c) Estimate the uncertainties Δp and Δx and then verify that $\Delta x \Delta p$ satisfies Heisenberg's uncertainty relation.

Exercise 1.40
Estimate the lifetime of the excited state of an atom whose natural width is 3×10^{-4} eV; you may need the value $\hbar = 6.626 \times 10^{-34}$ J s $= 4.14 \times 10^{-15}$ eV s.

Exercise 1.41
Calculate the final width of the wave packet corresponding to an 80 g bullet after traveling for 20 s; the size of the bullet is 2 cm.

Exercise 1.42
A 100 g arrow travels with a speed of 30 m s^{-1} over a distance of 50 m. If the initial size of the wave packet is 5 cm, what will be its final size?

Exercise 1.43
A 50 MeV beam of protons is fired over a distance of 10 km. If the initial size of the wave packet is 1.5×10^{-6} m, what will be the final size upon arrival?

Exercise 1.44
A 250 GeV beam of protons is fired over a distance of 1 km. If the initial size of the wave packet is 1 mm, find its final size.

Exercise 1.45
Consider an inextensible string of linear density μ (mass per unit length). If the string is subject to a tension T, the angular frequency of the string waves is given in terms of the wave number k by $\omega = k\sqrt{T/\mu}$. Find the phase and group velocities.

Exercise 1.46
The angular frequency for a wave propagating inside a waveguide is given in terms of the wave number k and the width b of the guide by $\omega = kc\left[1 - \pi^2/(b^2k^2)\right]^{-1/2}$. Find the phase and group velocities of the wave.

Exercise 1.47
Show that for those waves whose angular frequency ω and wave number k obey the dispersion relation $k^2c^2 = \omega^2 + constant$, the product of the phase and group velocities is equal to c^2, $v_g v_{ph} = c^2$, where c is the speed of light.

Exercise 1.48
How long will the wave packet of a 10 g object, initially confined to 1 mm, take to quadruple its size?

Exercise 1.49
How long will it take for the wave packet of a proton confined to 10^{-15} m to grow to a size equal to the distance between the Earth and the Sun? This distance is equal to 1.5×10^8 km.

Exercise 1.50
Assuming the wave packet representing the Moon to be confined to 1 m, how long will the packet take to reach a size triple that of the Sun? The Sun's radius is 6.96×10^5 km.

Chapter 2

Mathematical Tools of Quantum Mechanics

2.1 Introduction

We deal here with the mathematical machinery needed to study quantum mechanics. Although this chapter is mathematical in scope, no attempt is made to be mathematically complete or rigorous. We limit ourselves to those practical issues that are relevant to the formalism of quantum mechanics.

The Schrödinger equation is one of the cornerstones of the theory of quantum mechanics; it has the structure of a *linear* equation. The formalism of quantum mechanics deals with operators that are linear and wave functions that belong to an abstract Hilbert space. The mathematical properties and structure of Hilbert spaces are essential for a proper understanding of the formalism of quantum mechanics. For this, we are going to review briefly the properties of Hilbert spaces and those of linear operators. We will then consider Dirac's *bra-ket* notation.

Quantum mechanics was formulated in two different ways by Schrödinger and Heisenberg. Schrödinger's wave mechanics and Heisenberg's matrix mechanics are the representations of the general formalism of quantum mechanics in *continuous* and *discrete* basis systems, respectively. For this, we will also examine the mathematics involved in representing kets, bras, bra-kets, and operators in discrete and continuous bases.

2.2 The Hilbert Space and Wave Functions

2.2.1 The Linear Vector Space

A linear vector space consists of two sets of elements and two algebraic rules:

- a set of *vectors* ψ, ϕ, χ, ... and a set of *scalars* $a, b, c, \ldots$;
- a rule for vector *addition* and a rule for scalar *multiplication*.

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

(a) Addition rule
The addition rule has the properties and structure of an *abelian* group:

- If ψ and ϕ are vectors (elements) of a space, their sum, $\psi + \phi$, is also a vector of the same space.
- Commutativity: $\psi + \phi = \phi + \psi$.
- Associativity: $(\psi + \phi) + \chi = \psi + (\phi + \chi)$.
- Existence of a zero or neutral vector: for each vector ψ, there must exist a zero vector O such that: $O + \psi = \psi + O = \psi$.
- Existence of a symmetric or inverse vector: each vector ψ must have a symmetric vector $(-\psi)$ such that $\psi + (-\psi) = (-\psi) + \psi = O$.

(b) Multiplication rule
The multiplication of vectors by scalars (scalars can be real or complex numbers) has these properties:

- The product of a scalar with a vector gives another vector. In general, if ψ and ϕ are two vectors of the space, any linear combination $a\psi + b\phi$ is also a vector of the space, a and b being scalars.
- Distributivity with respect to addition:

$$a(\psi + \phi) = a\psi + a\phi, \qquad (a + b)\psi = a\psi + b\psi, \tag{2.1}$$

- Associativity with respect to multiplication of scalars:

$$a(b\psi) = (ab)\psi \tag{2.2}$$

- For each element ψ there must exist a unitary scalar I and a zero scalar "o" such that

$$I\psi = \psi I = \psi \quad \text{and} \quad o\psi = \psi o = o. \tag{2.3}$$

2.2.2 The Hilbert Space

A Hilbert space $\mathcal{H}$ consists of a set of vectors $\psi, \phi, \chi, \ldots$ and a set of *scalars* $a, b, c, \ldots$ which satisfy the following *four* properties:

(a) $\mathcal{H}$ is a linear space

The properties of a linear space were considered in the previous section.

(b) $\mathcal{H}$ has a defined scalar product that is strictly positive

The scalar product of an element ψ with another element ϕ is in general a complex number, denoted by (ψ, ϕ), where (ψ, ϕ) = complex number. **Note:** Watch out for the order! Since the scalar product is a complex number, the quantity (ψ, ϕ) is generally not equal to (ϕ, ψ): $(\psi, \phi) = \psi^* \phi$ while $(\phi, \psi) = \phi^* \psi$. The scalar product satisfies the following properties:

- The scalar product of ψ with ϕ is equal to the complex conjugate of the scalar product of ϕ with ψ:

$$(\psi, \phi) = (\phi, \psi)^*. \tag{2.4}$$

- The scalar product of ϕ with ψ is linear with respect to the second factor if $\psi = a\psi_1 + b\psi_2$:

$$(\phi, a\psi_1 + b\psi_2) = a(\phi, \psi_1) + b(\phi, \psi_2), \tag{2.5}$$

and antilinear with respect to the first factor if $\phi = a\phi_1 + b\phi_2$:

$$(a\phi_1 + b\phi_2, \psi) = a^*(\phi_1, \psi) + b^*(\phi_2, \psi). \tag{2.6}$$

- The scalar product of a vector ψ with itself is a positive real number:

$$(\psi, \psi) = \| \psi \|^2 \geq 0, \tag{2.7}$$

where the equality holds only for $\psi = O$.

(c) $\mathcal{H}$ is separable

There exists a Cauchy sequence $\psi_n \in \mathcal{H}$ $(n = 1, 2, \ldots)$ such that for every ψ of $\mathcal{H}$ and $\varepsilon > 0$, there exists at least one ψ_n of the sequence for which

$$\| \psi - \psi_n \| < \varepsilon. \tag{2.8}$$

(d) $\mathcal{H}$ is complete

Every Cauchy sequence $\psi_n \in \mathcal{H}$ converges to an element of $\mathcal{H}$. That is, for any ψ_n, the relation

$$\lim_{n,m \to \infty} \| \psi_n - \psi_m \| = 0, \tag{2.9}$$

defines a unique limit ψ of $\mathcal{H}$ such that

$$\lim_{n \to \infty} \| \psi - \psi_n \| = 0. \tag{2.10}$$

Remark
We should note that in a scalar product (ϕ, ψ), the second factor, ψ, belongs to the Hilbert space $\mathcal{H}$, while the first factor, ϕ, belongs to its dual Hilbert space $\mathcal{H}_d$. The distinction between $\mathcal{H}$ and $\mathcal{H}_d$ is due to the fact that, as mentioned above, the scalar product is not commutative: $(\phi, \psi) \neq (\psi, \phi)$; the order matters! From linear algebra, we know that every vector space can be associated with a dual vector space.

2.2.3 Dimension and Basis of a Vector Space

A set of N nonzero vectors $\phi_1, \phi_2, \ldots, \phi_N$ is said to be *linearly independent* if and only if the solution of the equation

$$\sum_{i=1}^{N} a_i \phi_i = 0 \tag{2.11}$$

is $a_1 = a_2 = \cdots = a_N = 0$. But if there exists a set of scalars, which are not all zero, so that one of the vectors (say ϕ_n) can be expressed as a linear combination of the others,

$$\phi_n = \sum_{i=1}^{n-1} a_i \phi_i + \sum_{i=n+1}^{N} a_i \phi_i, \tag{2.12}$$

the set $\{\phi_i\}$ is said to be *linearly dependent*.

Dimension: The *dimension* of a vector space is given by the *maximum number* of linearly independent vectors the space can have. For instance, if the maximum number of linearly independent vectors a space has is N (i.e., $\phi_1, \phi_2, \ldots, \phi_N$), this space is said to be N-dimensional. In this N-dimensional vector space, any vector ψ can be expanded as a linear combination:

$$\psi = \sum_{i=1}^{N} a_i \phi_i. \tag{2.13}$$

Basis: The *basis* of a vector space consists of a set of the maximum possible number of linearly independent vectors belonging to that space. This set of vectors, $\phi_1, \phi_2, \ldots, \phi_N$, to be denoted in short by $\{\phi_i\}$, is called the basis of the vector space, while the vectors $\phi_1, \phi_2, \ldots, \phi_N$ are called the base vectors. Although the set of these linearly independent vectors is arbitrary, it is convenient to choose them *orthonormal*; that is, their scalar products satisfy the relation $(\phi_i, \phi_j) = \delta_{ij}$ (we may recall that $\delta_{ij} = 1$ whenever $i = j$ and zero otherwise). The basis is said to be *orthonormal* if it consists of a set of orthonormal vectors. Moreover, the basis is said to be *complete* if it spans the entire space; that is, there is no need to introduce any additional base vector. The expansion coefficients a_i in (2.13) are called the *components* of the vector ψ in the basis. Each component is given by the scalar product of ψ with the corresponding base vector, $a_j = (\phi_j, \psi)$.

Examples of linear vector spaces

Let us give two examples of linear spaces that are Hilbert spaces: one having a *finite (discrete)* set of base vectors, the other an *infinite (continuous)* basis.

- The first one is the three-dimensional Euclidean vector space; the basis of a Cartesian coordinate system in this space consists of three linearly independent vectors, usually denoted by $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$. Any vector of the Euclidean space can be written in terms of the base vectors as $\vec{A} = a_1\hat{\imath} + a_2\hat{\jmath} + a_3\hat{k}$, where a_1, a_2, and a_3 are the components of $\vec{A}$ in the basis; each component can be determined by taking the scalar product of $\vec{A}$ with the corresponding base vector: $a_1 = \hat{\imath} \cdot \vec{A}$, $a_2 = \hat{\jmath} \cdot \vec{A}$, and $a_3 = \hat{k} \cdot \vec{A}$. Note that the scalar product in the Euclidean space is real and hence symmetric. The norm in this space is the usual length of vectors $\| \vec{A} \| = A$. Note also that whenever $a_1\hat{\imath} + a_2\hat{\jmath} + a_3\hat{k} = 0$ we have $a_1 = a_2 = a_3 = 0$ and that none of the unit vectors $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$ can be expressed as a linear combination of the other two.

- The second example is the space of the entire complex functions $\psi(x)$; the dimension of this space is infinite for it has an infinite number of linearly independent basis vectors.

Example 2.1
Check whether the following sets of functions are linearly independent or dependent on the real x-axis.

(a) $f(x) = 4, g(x) = x^2, h(x) = e^{2x}$
(b) $f(x) = x, g(x) = x^2, h(x) = x^3$
(c) $f(x) = x, g(x) = 5x, h(x) = x^2$
(d) $f(x) = 2 + x^2, g(x) = 3 - x + 4x^3, h(x) = 2x + 3x^2 - 8x^3$

Solution

(a) The first set is clearly linearly independent since $a_1 f(x) + a_2 g(x) + a_3 h(x) = 4a_1 + a_2 x^2 + a_3 e^{2x} = 0$ implies that $a_1 = a_2 = a_3 = 0$ for any value of x.

(b) The functions $f(x) = x, g(x) = x^2, h(x) = x^3$ are also linearly independent since $a_1 x + a_2 x^2 + a_3 x^3 = 0$ implies that $a_1 = a_2 = a_3 = 0$ no matter what the value of x. For instance, taking $x = -1, 1, 3$, the following system of three equations

$$-a_1 + a_2 - a_3 = 0, \qquad a_1 + a_2 + a_3 = 0, \qquad 3a_1 + 9a_2 + 27a_3 = 0 \tag{2.14}$$

yields $a_1 = a_2 = a_3 = 0$.

(c) The functions $f(x) = x, g(x) = 5x, h(x) = x^2$ are not linearly independent, since $g(x) = 5f(x) + 0 \times h(x)$.

(d) The functions $f(x) = 2 + x^2, g(x) = 3 - x + 4x^3, h(x) = 2x + 3x^2 - 8x^3$ are not linearly independent since $h(x) = 3f(x) - 2g(x)$.

Example 2.2
Are the following sets of vectors (in the three-dimensional Euclidean space) linearly independent or dependent?

(a) $\vec{A} = (3,0,0), \vec{B} = (0,-2,0), \vec{C} = (0,0,-1)$
(b) $\vec{A} = (6,-9,0), \vec{B} = (-2,3,0)$
(c) $\vec{A} = (2,3,-1), \vec{B} = (0,1,2), \vec{C} = (0,0,-5)$
(d) $\vec{A} = (1,-2,3), \vec{B} = (-4,1,7), \vec{C} = (0,10,11)$, and $\vec{D} = (14,3,-4)$

Solution

(a) The three vectors $\vec{A} = (3,0,0), \vec{B} = (0,-2,0), \vec{C} = (0,0,-1)$ are linearly independent, since

$$a_1\vec{A} + a_2\vec{B} + a_3\vec{C} = 0 \Longrightarrow 3a_1\hat{\imath} - 2a_2\hat{\jmath} - a_3\hat{k} = 0 \tag{2.15}$$

leads to

$$3a_1 = 0, \qquad -2a_2 = 0, \qquad -a_3 = 0, \tag{2.16}$$

which yields $a_1 = a_2 = a_3 = 0$.

(b) The vectors $\vec{A} = (6,-9,0), \vec{B} = (-2,3,0)$ are linearly dependent, since the solution to

$$a_1\vec{A} + a_2\vec{B} = 0 \quad \Longrightarrow \quad (6a_1 - 2a_2)\hat{\imath} + (-9a_1 + 3a_2)\hat{\jmath} = 0 \tag{2.17}$$

is $a_1 = a_2/3$. The first vector is equal to -3 times the second one: $\vec{A} = -3\vec{B}$.

(c) The vectors $\vec{A} = (2,3,-1), \vec{B} = (0,1,2), \vec{C} = (0,0,-5)$ are linearly independent, since

$$a_1\vec{A} + a_2\vec{B} + a_3\vec{C} = 0 \Longrightarrow 2a_1\hat{\imath} + (3a_1 + a_2)\hat{\jmath} + (-a_1 + 2a_2 - 5a_3)\hat{k} = 0 \tag{2.18}$$

leads to

$$2a_1 = 0, \qquad 3a_1 + a_2 = 0, \qquad -a_1 + 2a_2 - 5a_3 = 0. \tag{2.19}$$

The only solution of this system is $a_1 = a_2 = a_3 = 0$.

(d) The vectors $\vec{A} = (1,-2,3)$, $\vec{B} = (-4,1,7)$, $\vec{C} = (0,10,11)$, and $\vec{D} = (14,3,-4)$ are not linearly independent, because $\vec{D}$ can be expressed in terms of the other vectors:

$$\vec{D} = 2\vec{A} - 3\vec{B} + \vec{C}. \tag{2.20}$$

2.2.4 Square-Integrable Functions: Wave Functions

In the case of function spaces, a "vector" element is given by a *complex function* and the *scalar product* by *integrals*. That is, the scalar product of two functions $\psi(x)$ and $\phi(x)$ is given by

$$(\psi, \phi) = \int \psi^*(x)\phi(x)\,dx. \tag{2.21}$$

If this integral *diverges*, the scalar product *does not exist*. As a result, if we want the function space to possess a scalar product, we must select only those functions for which (ψ, ϕ) is *finite*. In particular, a function $\psi(x)$ is said to be *square integrable* if the scalar product of ψ with itself,

$$(\psi, \psi) = \int |\psi(x)|^2\,dx, \tag{2.22}$$

is *finite*.

It is easy to verify that the space of square-integrable functions possesses the properties of a Hilbert space. For instance, any linear combination of square-integrable functions is also a square-integrable function and (2.21) satisfies all the properties of the scalar product of a Hilbert space.

Note that the dimension of the Hilbert space of square-integrable functions is infinite, since each wave function can be expanded in terms of an infinite number of linearly independent functions. The dimension of a space is given by the maximum number of linearly independent basis vectors required to span that space.

A good example of square-integrable functions is the *wave function* of quantum mechanics, $\psi(\vec{r}, t)$. We have seen in Chapter 1 that, according to Born's probabilistic interpretation of $\psi(\vec{r}, t)$, the quantity $|\psi(\vec{r}, t)|^2\, d^3r$ represents the probability of finding, at time t, the particle in a volume d^3r, centered around the point $\vec{r}$. The probability of finding the particle somewhere in space must then be equal to 1:

$$\int |\psi(\vec{r}, t)|^2\, d^3r = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} |\psi(\vec{r}, t)|^2\, dz = 1; \tag{2.23}$$

hence the wave functions of quantum mechanics are square-integrable. Wave functions satisfying (2.23) are said to be normalized or square-integrable. As wave mechanics deals with square-integrable functions, any wave function which is not square-integrable has no physical meaning in quantum mechanics.

2.3 Dirac Notation

The physical state of a system is represented in quantum mechanics by elements of a Hilbert space; these elements are called state vectors; they are complex vectors that belong to an infinitely dimensional space. We can represent the state vectors in different bases by means of function expansions. This is analogous to specifying an ordinary (Euclidean) vector by its components in various coordinate systems. For instance, we can represent equivalently a vector by its components in a Cartesian coordinate system, in a spherical coordinate system, or in a cylindrical coordinate system. *The meaning of a vector is, of course, independent of the coordinate system chosen to represent its components.* Similarly, the state of a microscopic system has a meaning independent of the basis in which it is represented or expanded.

To free state vectors from coordinate meaning, Dirac introduced what was to become an invaluable notation in quantum mechanics; it allows one to manipulate the formalism of quantum mechanics with great ease and clarity. He introduced the concepts of kets, bras, and bra-kets, which will be explained below.

Kets: elements of a vector space
Dirac denoted the state vector ψ by the symbol $|\,\psi\rangle$, which he called a *ket* vector, or simply a ket; the symbol $|\,\rangle$ indicates that this is a vector and "ψ" is its label. Kets belong to the Hilbert (vector) space $\mathcal{H}$, or, in short, to the ket-space.

Bras: elements of a dual space
As mentioned above, we know from linear algebra that a dual space can be associated with every vector space. Dirac denoted the elements of a dual space by the symbol $\langle\,|$, which he called a bra vector, or simply a bra; for instance, the element $\langle\psi\,|$ represents a bra. *Note: For every ket* $|\,\psi\rangle$ *there exists a unique bra* $\langle\psi\,|$ *and vice versa.* Again, while kets belong to the Hilbert space $\mathcal{H}$, the corresponding bras belong to its dual (Hilbert) space $\mathcal{H}_d$. In Section 2.5.1, we will see that kets can be represented by column vectors and bras by row vectors.

Bra-ket: Dirac notation for the scalar product
Dirac denoted the scalar (inner) product by the symbol $\langle\,|\,\rangle$, which he called a a *bra-ket*. For instance, the scalar product (ϕ, ψ) is denoted by the bra-ket $\langle\phi\,|\,\psi\rangle$:

$$(\phi, \psi) \quad \longrightarrow \quad \langle\phi\,|\,\psi\rangle. \tag{2.24}$$

Note: When a ket (or bra) is multiplied by a complex number, we also get a ket (or bra).

Remark: In wave mechanics we deal with wave functions $\psi(\vec{r}, t)$, but in the more general formalism of quantum mechanics we deal with abstract kets $|\,\psi\rangle$. Wave functions, like kets, are elements of a Hilbert space. We should note that, like a wave function, a ket represents the system completely, and hence knowing $|\,\psi\rangle$ means knowing all its amplitudes in all possible representations. As mentioned above, kets are independent of any particular representation. There is no reason to single out a particular representation basis such as the representation in the position space. Of course, if we want to know the probability of finding the particle at some position in space, we need to work out the formalism within the coordinate representation. The state vector of this particle at time t will be given by the spatial wave function $\langle\vec{r}, t\,|\,\psi\rangle = \psi(\vec{r}, t)$. In the coordinate representation, the scalar product $\langle\phi\,|\,\psi\rangle$ is given by

$$\langle\phi\,|\,\psi\rangle = \int \phi^*(\vec{r}, t)\psi(\vec{r}, t)\, d^3r. \tag{2.25}$$

Similarly, if we are considering the three-dimensional momentum of a particle, the ket $|\psi\rangle$ will have to be expressed in momentum space. In this case the state of the particle will be described by a wave function $\psi(\vec{p}, t)$, where $\vec{p}$ is the momentum of the particle.

Properties of kets, bras, and bra-kets

- **Every ket has a corresponding bra**

  To every *ket* $|\psi\rangle$, there corresponds a unique *bra* $\langle\psi|$ and vice versa:

$$|\psi\rangle \quad \longleftrightarrow \quad \langle\psi| \,. \tag{2.26}$$

  There is a one-to-one correspondence between bras and kets:

$$a|\psi\rangle + b|\phi\rangle \quad \longleftrightarrow \quad a^*\langle\psi| + b^*\langle\phi|, \tag{2.27}$$

  where a and b are complex numbers.

- **Properties of the scalar product**

  In quantum mechanics, since the scalar product is a complex number, the ordering matters a lot. We must be careful to distinguish a scalar product from its complex conjugate; $\langle\psi|\phi\rangle$ is not the same thing as $\langle\phi|\psi\rangle$:

$$\langle\phi|\psi\rangle^* = \langle\psi|\phi\rangle. \tag{2.28}$$

  This property becomes clearer if we apply it to (2.21):

$$\langle\phi|\psi\rangle^* = \left(\int \phi^*(\vec{r}, t)\psi(\vec{r}, t)\, d^3r\right)^* = \int \psi^*(\vec{r}, t)\phi(\vec{r}, t)\, d^3r = \langle\psi|\phi\rangle. \tag{2.29}$$

  When $|\psi\rangle$ and $|\phi\rangle$ are real, we would have $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle$. Let us list some additional properties of the scalar product:

$$\langle\psi|(a_1|\psi_1\rangle + a_2|\psi_2\rangle) = a_1\langle\psi|\psi_1\rangle + a_2\langle\psi|\psi_2\rangle, \tag{2.30}$$

$$\left(a_1\langle\phi_1| + a_2\langle\phi_2|\right)|\psi\rangle = a_1\langle\phi_1|\psi\rangle + a_2\langle\phi_2|\psi\rangle, \tag{2.31}$$

$$\left(a_1\langle\phi_1| + a_2\langle\phi_2|\right)(b_1|\psi_1\rangle + b_2|\psi_2\rangle) = a_1b_1\langle\phi_1|\psi_1\rangle + a_1b_2\langle\phi_1|\psi_2\rangle + a_2b_1\langle\phi_2|\psi_1\rangle + a_2b_2\langle\phi_2|\psi_2\rangle. \tag{2.32}$$

- **The norm is real and positive**

  For any state vector $|\psi\rangle$ of the Hilbert space $\mathcal{H}$, the norm $\langle\psi|\psi\rangle$ is real and positive; $\langle\psi|\psi\rangle$ is equal to zero only for the case where $|\psi\rangle = O$, where O is the zero vector. If the state $|\psi\rangle$ is normalized then $\langle\psi|\psi\rangle = 1$.

- **Schwarz inequality**

  For any two states $|\psi\rangle$ and $|\phi\rangle$ of the Hilbert space, we can show that

$$\left|\langle\psi|\phi\rangle\right|^2 \leq \langle\psi|\psi\rangle\langle\phi|\phi\rangle. \tag{2.33}$$

If $|\psi\rangle$ and $|\phi\rangle$ are linearly dependent (i.e., proportional: $|\psi\rangle = \alpha\,|\phi\rangle$, where α is a scalar), this relation becomes an equality. The Schwarz inequality (2.33) is analogous to the following relation of the real Euclidean space

$$|\vec{A}\cdot\vec{B}|^2 \leq |\vec{A}|^2\,|\vec{B}|^2\,. \tag{2.34}$$

- **Triangle inequality**

$$\sqrt{\langle\psi+\phi\,|\,\psi+\phi\rangle} \leq \sqrt{\langle\psi\,|\,\psi\rangle} + \sqrt{\langle\phi\,|\,\phi\rangle}. \tag{2.35}$$

If $|\psi\rangle$ and $|\phi\rangle$ are linearly dependent, $|\psi\rangle = \alpha\,|\phi\rangle$, and if the proportionality scalar α is real and positive, the triangle inequality becomes an equality. The counterpart of this inequality in Euclidean space is given by $|\vec{A}+\vec{B}| \leq |\vec{A}| + |\vec{B}|$.

- **Orthogonal states**

Two *kets*, $|\psi\rangle$ and $|\phi\rangle$, are said to be orthogonal if they have a vanishing scalar product:

$$\langle\psi\,|\,\phi\rangle = 0. \tag{2.36}$$

- **Orthonormal states**

Two *kets*, $|\psi\rangle$ and $|\phi\rangle$, are said to be orthonormal if they are orthogonal and if each one of them has a unit norm:

$$\langle\psi\,|\,\phi\rangle = 0, \qquad \langle\psi\,|\,\psi\rangle = 1, \qquad \langle\phi\,|\,\phi\rangle = 1. \tag{2.37}$$

- **Forbidden quantities**

If $|\psi\rangle$ and $|\phi\rangle$ belong to the same vector (Hilbert) space, products of the type $|\psi\rangle\,|\phi\rangle$ and $\langle\psi\,|\,\langle\phi\,|$ are forbidden. They are nonsensical, since $|\psi\rangle\,|\phi\rangle$ and $\langle\psi\,|\,\langle\phi\,|$ are neither kets nor bras (an explicit illustration of this will be carried out in the example below and later on when we discuss the representation in a discrete basis). If $|\psi\rangle$ and $|\phi\rangle$ belong, however, to different vector spaces (e.g., $|\psi\rangle$ belongs to a spin space and $|\phi\rangle$ to an orbital angular momentum space), then the product $|\psi\rangle\,|\phi\rangle$, written as $|\psi\rangle\otimes|\phi\rangle$, represents a tensor product of $|\psi\rangle$ and $|\phi\rangle$. Only in these typical cases are such products meaningful.

- **Important Note about the use of the notations like $|a\psi\rangle$, $\langle b\psi|$, and $|\psi+\phi\rangle$:**
First, the two notations $|a\psi\rangle$ and $\langle b\psi|$ and their respective relations to $a\,|\psi\rangle$ and $b^*\langle\psi|$ are often encountered in the literature (including the second edition of this book); they are often equated as follows: $a\,|\psi\rangle = |a\psi\rangle$ and $b^*\langle\psi| = \langle b\psi|$. However, one has to be careful not to commit the following error when numbers are taken inside a ket or a bra: the notations $|a\psi\rangle$ and $\langle b\psi|$ do not mean that we multiply the label "ψ" with the numbers "a" and "b" inside the ket and bra, respectively. For instance, if we are dealing with a basis consisting of three unit vectors $|1\rangle$, $|2\rangle$, $|3\rangle$, it will be wrong to write $2\,|1\rangle = |2\rangle$, because the notation $2\,|1\rangle$ simply means that we need to multiply each one of the components of the basis vector $|1\rangle$ by the number 2 and that the result is certainly

different from the unit vector $|\,2\rangle$; that is, $2\,|\,1\rangle \neq |\,2\rangle$. As such, in order to prevent this kind of error from occurring, it will be safer to write:

$$a\,|\,\psi\rangle \neq |\,a\psi\rangle, \qquad b^*\langle\psi\,| \neq \langle b\psi\,|\,. \tag{2.38}$$

Due to the confusion and errors that can potentially result from the misuse of notations like $|\,a\psi\rangle$ and $\langle b\psi\,|$, and as illustrated below in Example 2.4 below, I have avoided using them altogether throughout this edition.

Second, one has to be careful also with notations like $|\,\psi+\phi\rangle$ and $\langle\psi+\phi\,|$ where one should never add the labels "ψ" and "ϕ" inside a ket or bra; the correct way to add these expressions is as follows:

$$|\,\psi+\phi\rangle = |\,\psi\rangle + |\,\phi\rangle, \qquad \langle\psi+\phi\,| = \langle\psi\,| + \langle\phi\,|\,. \tag{2.39}$$

For instance, one should write $|\,1+2\rangle$ like $|\,1+2\rangle = |\,1\rangle + |\,2\rangle$, but never like $|\,1+2\rangle = |\,3\rangle$, since $|\,1+2\rangle \neq |\,3\rangle$. See also Example 2.4 for a more explicit illustration.

Example 2.3

(**Note:** We will see later in this chapter (see Section 2.5.1) that kets are represented by column matrices, bras by row matrices, and brakets by complex numbers; this example is offered earlier than it should because we felt the need to show some concrete illustrations of the formalism at this very level.) Consider the following two kets:

$$|\,\psi\rangle = \begin{pmatrix} -3i \\ 2+i \\ 4 \end{pmatrix}, \qquad |\,\phi\rangle = \begin{pmatrix} 2 \\ -i \\ 2-3i \end{pmatrix}.$$

(a) Find the bra $\langle\phi\,|$.

(b) Evaluate the scalar product $\langle\phi\,|\,\psi\rangle$.

(c) Examine why the products $|\,\psi\rangle\,|\,\phi\rangle$ and $\langle\phi\,|\,\langle\psi\,|$ do not make sense.

Solution

(a) As will be explained later when we introduce the Hermitian adjoint of kets and bras, we want to mention that the bra $\langle\phi\,|$ can be obtained by simply taking the complex conjugate of the transpose of the ket $|\,\phi\rangle$:

$$\langle\phi\,| = (2 \quad i \quad 2+3i)\,. \tag{2.40}$$

(b) The scalar product $\langle\phi\,|\,\psi\rangle$ can be calculated as follows:

$$\begin{aligned} \langle\phi\,|\,\psi\rangle &= (2 \quad i \quad 2+3i)\begin{pmatrix} -3i \\ 2+i \\ 4 \end{pmatrix} \\ &= 2(-3i) + i(2+i) + 4(2+3i) \\ &= 7+8i. \end{aligned} \tag{2.41}$$

(c) First, the product $|\,\psi\rangle\,|\,\phi\rangle$ cannot be performed because, from linear algebra, the product of two column matrices cannot be performed. Similarly, since two row matrices cannot be multiplied, the product $\langle\phi\,|\,\langle\psi\,|$ is meaningless.

Example 2.4

(Note: This example is designed to illustrate the type of errors that can occur when using notations like $|a\psi\rangle$ and $|\psi+\phi\rangle$ where one might unwittingly multiply the label ψ" by "a" in the first ket or add the labels "ψ" and "ϕ" in the second ket.)

Consider a space which is spanned by three basis vectors

$$|1\rangle = \frac{1}{5}\begin{pmatrix} 4+3i \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad |2\rangle = \frac{1}{2}\begin{pmatrix} 0 \\ 1+i \\ 0 \\ 0 \end{pmatrix}, \qquad |3\rangle = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ 1-i \\ 0 \end{pmatrix}, \qquad |4\rangle = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ \\ 0 \\ \sqrt{3}+i \end{pmatrix}.$$

(a) Find the ket $3\,|\,1\rangle$. Is it equal to $|\,3\rangle$?
(b) Find the ket $2\,|\,2\rangle$. Is it equal to $|\,4\rangle$?
(c) Find the ket $|\,1+3\rangle$. Is it equal to $|\,4\rangle$?

Solution

(a) The ket $3\,|\,1\rangle$ can be obtained by simply multiplying all the components of $|\,1\rangle$ by 3:

$$3\,|\,1\rangle = \frac{3}{5}\begin{pmatrix} 4+3i \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 12+9i \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{2.42}$$

Clearly $3\,|\,1\rangle$ has nothing to do with the ket $|\,3\rangle$:

$$3\,|\,1\rangle \neq |\,3\rangle, \tag{2.43}$$

since

$$3\,|\,1\rangle = \frac{1}{5}\begin{pmatrix} 12+9i \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \text{whereas} \qquad |\,3\rangle = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ 1-i \\ 0 \end{pmatrix}. \tag{2.44}$$

(b) Similarly, the ket $2\,|\,2\rangle$ can be obtained by simply multiplying all the components of $|\,2\rangle$ by 2:

$$2\,|\,2\rangle = \frac{2}{2}\begin{pmatrix} 0 \\ 1+i \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1+i \\ 0 \\ 0 \end{pmatrix}. \tag{2.45}$$

Again, $2\,|\,2\rangle$ has nothing to do with the ket $|\,4\rangle$:

$$2\,|\,2\rangle \neq |\,4\rangle, \tag{2.46}$$

since

$$2\,|\,2\rangle = \begin{pmatrix} 0 \\ 1+i \\ 0 \\ 0 \end{pmatrix}, \qquad \text{whereas} \qquad |\,4\rangle = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ \\ 0 \\ \sqrt{3}+i \end{pmatrix}. \tag{2.47}$$

(c) The ket $|1+3\rangle$ is obtained by simply adding the components of $|1\rangle$ to those of $|3\rangle$:

$$|1+3\rangle = |1\rangle + |3\rangle = \frac{1}{5}\begin{pmatrix}4+3i\\0\\0\\0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0\\0\\1-i\\0\end{pmatrix} = \begin{pmatrix}(4+3i)/5\\0\\(1-i)/2\\0\end{pmatrix}. \tag{2.48}$$

Clearly, $|1+3\rangle$ has nothing to do with the ket $|4\rangle$:

$$|1+3\rangle \neq |4\rangle, \tag{2.49}$$

since

$$|1+3\rangle = \begin{pmatrix}(4+3i)/5\\0\\(1-i)/2\\0\end{pmatrix}, \qquad \text{whereas} \qquad |4\rangle = \frac{1}{2}\begin{pmatrix}0\\0\\0\\\sqrt{3}+i\end{pmatrix}. \tag{2.50}$$

Physical meaning of the scalar product

The scalar product can be interpreted in two ways. First, by analogy with the scalar product of ordinary vectors in the Euclidean space, where $\vec{A}\cdot\vec{B}$ represents the projection of $\vec{B}$ on $\vec{A}$, the product $\langle\phi \mid \psi\rangle$ also represents the projection of $|\psi\rangle$ onto $|\phi\rangle$. Second, in the case of normalized states and according to Born's probabilistic interpretation, the quantity $\langle\phi \mid \psi\rangle$ represents the probability amplitude that the system's state $|\psi\rangle$ will, after a measurement is performed on the system, be found to be in another state $|\phi\rangle$.

Example 2.5 (Bra-ket algebra)

Consider the states $|\psi\rangle = 3i\,|\phi_1\rangle - 7i\,|\phi_2\rangle$ and $|\chi\rangle = -\,|\phi_1\rangle + 2i\,|\phi_2\rangle$, where $|\phi_1\rangle$ and $|\phi_2\rangle$ are orthonormal.

(a) Calculate $|\psi+\chi\rangle$ and $\langle\psi+\chi|$.

(b) Calculate the scalar products $\langle\psi \mid \chi\rangle$ and $\langle\chi \mid \psi\rangle$. Are they equal?

(c) Show that the states $|\psi\rangle$ and $|\chi\rangle$ satisfy the Schwarz inequality.

(d) Show that the states $|\psi\rangle$ and $|\chi\rangle$ satisfy the triangle inequality.

Solution

(a) The calculation of $|\psi+\chi\rangle$ is straightforward:

$$\begin{aligned}|\psi+\chi\rangle &= |\psi\rangle + |\chi\rangle = \big(3i\,|\phi_1\rangle - 7i\,|\phi_2\rangle\big) + \big(-\,|\phi_1\rangle + 2i\,|\phi_2\rangle\big)\\ &= (-1+3i)\,|\phi_1\rangle - 5i\,|\phi_2\rangle.\end{aligned} \tag{2.51}$$

This leads at once to the expression of $\langle\psi+\chi|$:

$$\langle\psi+\chi| = (-1+3i)^*\langle\phi_1| + (-5i)^*\langle\phi_2| = (-1-3i)\langle\phi_1| + 5i\langle\phi_2|\,. \tag{2.52}$$

(b) Since $\langle\phi_1 \mid \phi_1\rangle = \langle\phi_2 \mid \phi_2\rangle = 1$, $\langle\phi_1 \mid \phi_2\rangle = \langle\phi_2 \mid \phi_1\rangle = 0$, and since the bras corresponding to the kets $|\psi\rangle = 3i\,|\phi_1\rangle - 7i\,|\phi_2\rangle$ and $|\chi\rangle = -\,|\phi_1\rangle + 2i\,|\phi_2\rangle$ are given by

$\langle\psi| = -3i\langle\phi_1| + 7i\langle\phi_2|$ and $\langle\chi| = -\langle\phi_1| - 2i\langle\phi_2|$, the scalar products are

$$\begin{aligned}\langle\psi\,|\,\chi\rangle &= \big(-3i\langle\phi_1| + 7i\langle\phi_2|\big)\big(-|\,\phi_1\rangle + 2i\,|\,\phi_2\rangle\big)\\ &= (-3i)(-1)\langle\phi_1\,|\,\phi_1\rangle + (7i)(2i)\langle\phi_2\,|\,\phi_2\rangle\\ &= -14 + 3i, \qquad (2.53)\\ \langle\chi\,|\,\psi\rangle &= \big(-\langle\phi_1| - 2i\langle\phi_2|\big)\big(3i\,|\,\phi_1\rangle - 7i\,|\,\phi_2\rangle\big)\\ &= (-1)(3i)\langle\phi_1\,|\,\phi_1\rangle + (-2i)(-7i)\langle\phi_2\,|\,\phi_2\rangle\\ &= -14 - 3i. \qquad (2.54)\end{aligned}$$

We see that $\langle\psi\,|\,\chi\rangle$ is equal to the complex conjugate of $\langle\chi\,|\,\psi\rangle$.

(c) Let us first calculate $\langle\psi\,|\,\psi\rangle$ and $\langle\chi\,|\,\chi\rangle$:

$$\langle\psi\,|\,\psi\rangle = \big(-3i\langle\phi_1| + 7i\langle\phi_2|\big)\big(3i\,|\,\phi_1\rangle - 7i\,|\,\phi_2\rangle\big) = (-3i)(3i) + (7i)(-7i) = 58, \qquad (2.55)$$

$$\langle\chi\,|\,\chi\rangle = \big(-\langle\phi_1| - 2i\langle\phi_2|\big)\big(-|\,\phi_1\rangle + 2i\,|\,\phi_2\rangle\big) = (-1)(-1) + (-2i)(2i) = 5. \qquad (2.56)$$

Since $\langle\psi\,|\,\chi\rangle = -14+3i$ we have $|\,\langle\psi\,|\,\chi\rangle\,|^2 = 14^2+3^2 = 205$. Combining the values of $|\,\langle\psi\,|\,\chi\rangle\,|^2$, $\langle\psi\,|\,\psi\rangle$, and $\langle\chi\,|\,\chi\rangle$, we see that the Schwarz inequality (2.33) is satisfied:

$$205 < (58)(5) \Longrightarrow |\,\langle\psi\,|\,\chi\rangle\,|^2 < \langle\psi\,|\,\psi\rangle\langle\chi\,|\,\chi\rangle. \qquad (2.57)$$

(d) First, let us use (2.51) and (2.52) to calculate $\langle\psi+\chi\,|\,\psi+\chi\rangle$:

$$\begin{aligned}\langle\psi+\chi\,|\,\psi+\chi\rangle &= \big[(-1-3i)\langle\phi_1| + 5i\langle\phi_2|\big]\big[(-1+3i)\,|\,\phi_1\rangle - 5i\,|\,\phi_2\rangle\big]\\ &= (-1-3i)(-1+3i) + (5i)(-5i)\\ &= 35. \qquad (2.58)\end{aligned}$$

Since $\langle\psi\,|\,\psi\rangle = 58$ and $\langle\chi\,|\,\chi\rangle = 5$, we infer that the triangle inequality (2.35) is satisfied:

$$\sqrt{35} < \sqrt{58} + \sqrt{5} \Longrightarrow \sqrt{\langle\psi+\chi\,|\,\psi+\chi\rangle} < \sqrt{\langle\psi\,|\,\psi\rangle} + \sqrt{\langle\chi\,|\,\chi\rangle}. \qquad (2.59)$$

Example 2.6

Consider two states $|\psi_1\rangle = 2i|\phi_1\rangle + |\phi_2\rangle - a|\phi_3\rangle + 4|\phi_4\rangle$ and $|\psi_2\rangle = 3|\phi_1\rangle - i|\phi_2\rangle + 5|\phi_3\rangle - |\phi_4\rangle$, where $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$, and $|\phi_4\rangle$ are orthonormal kets, and where a is a constant. Find the value of a so that $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthogonal.

Solution

For the states $|\psi_1\rangle$ and $|\psi_2\rangle$ to be orthogonal, the scalar product $\langle\psi_2\,|\,\psi_1\rangle$ must be zero. Using the relation $\langle\psi_2| = 3\langle\phi_1| + i\langle\phi_2| + 5\langle\phi_3| - \langle\phi_4|$, we can easily find the scalar product

$$\begin{aligned}\langle\psi_2\,|\,\psi_1\rangle &= \big(3\langle\phi_1| + i\langle\phi_2| + 5\langle\phi_3| - \langle\phi_4|\big)\big(2i|\phi_1\rangle + |\phi_2\rangle - a|\phi_3\rangle + 4|\phi_4\rangle\big)\\ &= 7i - 5a - 4. \qquad (2.60)\end{aligned}$$

Since $\langle\psi_2\,|\,\psi_1\rangle = 7i - 5a - 4 = 0$, the value of a is $a = (7i-4)/5$.

2.4 Operators

2.4.1 General Definitions

Definition of an operator: An operator[1] $\hat{A}$ is a *mathematical rule* that when applied to a ket $|\psi\rangle$ transforms it into another ket $|\psi'\rangle$ of the same space and when it acts on a bra $\langle\phi|$ transforms it into another bra $\langle\phi'|$:

$$\hat{A}|\psi\rangle = |\psi'\rangle, \qquad \langle\phi|\hat{A} = \langle\phi'|. \tag{2.61}$$

A similar definition applies to wave functions:

$$\hat{A}\psi(\vec{r}) = \psi'(\vec{r}), \qquad \phi(\vec{r})\hat{A} = \phi'(\vec{r}). \tag{2.62}$$

Examples of operators
Here are some of the operators that we will use in this text:

- Unity operator: it leaves any ket unchanged, $\hat{I}|\psi\rangle = |\psi\rangle$.
- The gradient operator: $\vec{\nabla}\psi(\vec{r}) = (\partial\psi(\vec{r})/\partial x)\hat{\imath} + (\partial\psi(\vec{r})/\partial y)\hat{\jmath} + (\partial\psi(\vec{r})/\partial z)\hat{k}$.
- The linear momentum operator: $\vec{P}\psi(\vec{r}) = -i\hbar\vec{\nabla}\psi(\vec{r})$.
- The Laplacian operator: $\nabla^2\psi(\vec{r}) = \partial^2\psi(\vec{r})/\partial x^2 + \partial^2\psi(\vec{r})/\partial y^2 + \partial^2\psi(\vec{r})/\partial z^2$.
- The parity operator: $\hat{\mathcal{P}}\psi(\vec{r}) = \psi(-\vec{r})$.

Remarks

- Operators act only on kets and bras; they don't act on scalars. For instance, for an operator $\hat{A}$ and a scalar b, we have: $\hat{A}b = b\hat{A}$ where $b\hat{A}$ is an operator.
- The integer power of an operator gives an operator: $\hat{A}^n$ = operator.

Products of operators
The product of two operators is generally not commutative:

$$\hat{A}\hat{B} \neq \hat{B}\hat{A}. \tag{2.63}$$

The product of operators is, however, associative:

$$\hat{A}\hat{B}\hat{C} = \hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C}. \tag{2.64}$$

We may also write $\hat{A}^n\hat{A}^m = \hat{A}^{n+m}$. When the product $\hat{A}\hat{B}$ operates on a ket $|\psi\rangle$ (the order of application is important), the operator $\hat{B}$ acts first on $|\psi\rangle$ and then $\hat{A}$ acts on the new ket $(B|\psi\rangle)$:

$$\hat{A}\hat{B}|\psi\rangle = \hat{A}(\hat{B}|\psi\rangle). \tag{2.65}$$

[1] The hat on $\hat{A}$ will be used throughout this text to distinguish an operator $\hat{A}$ from a complex number or a matrix A.

Similarly, when $\hat{A}\hat{B}\hat{C}\hat{D}$ operates on a ket $|\psi\rangle$, $\hat{D}$ acts first, then $\hat{C}$, then $\hat{B}$, and then $\hat{A}$.

When an operator $\hat{A}$ is sandwiched between a bra $\langle\phi|$ and a ket $|\psi\rangle$, it yields in general a complex number: $\langle\phi|\hat{A}|\psi\rangle$ = complex number. The quantity $\langle\phi|\hat{A}|\psi\rangle$ can also be a purely real or a purely imaginary number. **Note:** In evaluating $\langle\phi|\hat{A}|\psi\rangle$ it does not matter if one first applies $\hat{A}$ to the ket and then takes the bra-ket or one first applies $\hat{A}$ to the bra and then takes the bra-ket; that is $(\langle\phi|\hat{A})|\psi\rangle = \langle\phi|(\hat{A}|\psi\rangle)$.

Linear operators

An operator $\hat{A}$ is said to be *linear* if it obeys the distributive law and, like all operators, it commutes with constants. That is, an operator $\hat{A}$ is linear if, for any vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ and any complex numbers a_1 and a_2, we have

$$\hat{A}(a_1|\psi_1\rangle + a_2|\psi_2\rangle) = a_1\hat{A}|\psi_1\rangle + a_2\hat{A}|\psi_2\rangle, \tag{2.66}$$

and

$$(\langle\psi_1|a_1 + \langle\psi_2|a_2)\hat{A} = a_1\langle\psi_1|\hat{A} + a_2\langle\psi_2|\hat{A}. \tag{2.67}$$

Remarks

- The *expectation* or *mean* value $\langle\hat{A}\rangle$ of an operator $\hat{A}$ with respect to a state $|\psi\rangle$ is defined by

$$\langle\hat{A}\rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}. \tag{2.68}$$

- The quantity $|\phi\rangle\langle\psi|$ (i.e., the product of a ket with a bra) is a linear operator in Dirac's notation. To see this, when $|\phi\rangle\langle\psi|$ is applied to a ket $|\psi'\rangle$, we obtain another ket:

$$|\phi\rangle\langle\psi|\psi'\rangle = \langle\psi|\psi'\rangle|\phi\rangle, \tag{2.69}$$

  since $\langle\psi|\psi'\rangle$ is a complex number.

- We can show the following relation:

$$\left(|\phi\rangle\langle\psi|\right)^n = \langle\psi|\phi\rangle^{n-1}|\phi\rangle\langle\psi|, \tag{2.70}$$

  where $\langle\psi|\phi\rangle^{n-1}$ is a scalar and $|\phi\rangle\langle\psi|$ is an operator.

- **Forbidden quantities:** Products of the type $|\psi\rangle\hat{A}$ and $\hat{A}\langle\psi|$ (i.e., when an operator stands on the right of a ket or on the left of a bra) are forbidden. They are not operators, or kets, or bras; they have no mathematical or physical meanings (see equation (2.253) for an illustration).

2.4.2 Hermitian Adjoint

The Hermitian adjoint or conjugate[2], $\alpha^\dagger$, of a complex number α is the complex conjugate of this number: $\alpha^\dagger = \alpha^*$. The Hermitian adjoint, or simply the adjoint, $\hat{A}^\dagger$, of an operator $\hat{A}$ is defined by this relation:

$$\langle\psi|\hat{A}^\dagger|\phi\rangle = \langle\phi|\hat{A}|\psi\rangle^*. \tag{2.71}$$

[2]The terms "adjoint" and "conjugate" are used indiscriminately.

Properties of the Hermitian conjugate rule

To obtain the Hermitian adjoint of any expression, we must cyclically reverse the order of the factors and make three replacements:

- Replace constants by their complex conjugates: $\alpha^\dagger = \alpha^*$.
- Replace kets (bras) by the corresponding bras (kets): $(|\psi\rangle)^\dagger = \langle\psi|$ and $(\langle\psi|)^\dagger = |\psi\rangle$.
- Replace operators by their adjoints.

Following these rules, we can write

$$(\hat{A}^\dagger)^\dagger = \hat{A}, \tag{2.72}$$

$$(a\hat{A})^\dagger = a^*\hat{A}^\dagger, \tag{2.73}$$

$$(\hat{A}^n)^\dagger = (\hat{A}^\dagger)^n, \tag{2.74}$$

$$(\hat{A}+\hat{B}+\hat{C}+\hat{D})^\dagger = \hat{A}^\dagger+\hat{B}^\dagger+\hat{C}^\dagger+\hat{D}^\dagger, \tag{2.75}$$

$$(\hat{A}\hat{B}\hat{C}\hat{D})^\dagger = \hat{D}^\dagger\hat{C}^\dagger\hat{B}^\dagger\hat{A}^\dagger, \tag{2.76}$$

$$(\hat{A}\hat{B}\hat{C}\hat{D}\,|\,\psi\rangle)^\dagger = \langle\psi\,|\,D^\dagger C^\dagger B^\dagger A^\dagger. \tag{2.77}$$

The Hermitian adjoint of the operator $|\psi\rangle\langle\phi|$ is given by

$$(|\psi\rangle\langle\phi|)^\dagger = |\phi\rangle\langle\psi|. \tag{2.78}$$

Operators act inside kets and bras, respectively, as follows:

$$|\alpha\hat{A}\psi\rangle = \alpha\hat{A}\,|\psi\rangle, \qquad \langle\alpha\hat{A}\psi| = \alpha^*\langle\psi|\hat{A}^\dagger. \tag{2.79}$$

Note also that $\langle\alpha\hat{A}^\dagger\psi| = \alpha^*\langle\psi|(\hat{A}^\dagger)^\dagger = \alpha^*\langle\psi|\hat{A}$. Hence, we can also write:

$$\langle\psi|\hat{A}|\phi\rangle = \langle\hat{A}^\dagger\psi|\phi\rangle = \langle\psi|\hat{A}\phi\rangle. \tag{2.80}$$

Hermitian and skew-Hermitian operators

An operator $\hat{A}$ is said to be *Hermitian* if it is equal to its adjoint $\hat{A}^\dagger$:

$$\boxed{\text{Hermitian Operator:} \quad \hat{A} = \hat{A}^\dagger \quad \text{or} \quad \langle\psi|\hat{A}|\phi\rangle = \langle\phi|\hat{A}|\psi\rangle^*.} \tag{2.81}$$

On the other hand, an operator $\hat{B}$ is said to be *skew-Hermitian* or *anti-Hermitian* if

$$\boxed{\text{Anti-Hermitian Operator:} \quad \hat{B}^\dagger = -\hat{B} \quad \text{or} \quad \langle\psi|\hat{B}|\phi\rangle = -\langle\phi|\hat{B}|\psi\rangle^*.} \tag{2.82}$$

Remark

The Hermitian adjoint of an operator is not, in general, equal to its complex conjugate: $\hat{A}^\dagger \neq \hat{A}^*$.

Example 2.7

(a) Discuss the hermiticity of the operators $(\hat{A} + \hat{A}^\dagger)$, $i(\hat{A} + \hat{A}^\dagger)$, and $i(\hat{A} - \hat{A}^\dagger)$.

(b) Find the Hermitian adjoint of $f(\hat{A}) = (1 + i\hat{A} + 3\hat{A}^2)(1 - 2i\hat{A} - 9\hat{A}^2)/(5 + 7\hat{A})$.

(c) Show that the expectation value of a Hermitian operator is real and that of an anti-Hermitian operator is imaginary.

Solution

(a) The operator $\hat{B} = \hat{A} + \hat{A}^\dagger$ is Hermitian regardless of whether or not $\hat{A}$ is Hermitian, since

$$\hat{B}^\dagger = (\hat{A} + \hat{A}^\dagger)^\dagger = \hat{A}^\dagger + \hat{A} = \hat{B}. \tag{2.83}$$

Similarly, the operator $i(\hat{A} - \hat{A}^\dagger)$ is also Hermitian, but $i(\hat{A} + \hat{A}^\dagger)$ is anti-Hermitian, since

$$\left[i\left(\hat{A} + \hat{A}^\dagger\right)\right]^\dagger = -i\left(\hat{A}^\dagger + \hat{A}\right) = -i\left(\hat{A} + \hat{A}^\dagger\right), \tag{2.84}$$

$$\left[i\left(\hat{A} - \hat{A}^\dagger\right)\right]^\dagger = -i\left(\hat{A}^\dagger - \hat{A}\right) = i\left(\hat{A} - \hat{A}^\dagger\right). \tag{2.85}$$

(b) Since the Hermitian adjoint of an operator function $f(\hat{A})$ is given by $f^\dagger(\hat{A}) = f^*(\hat{A}^\dagger)$, we can write

$$\left(\frac{(1 + i\hat{A} + 3\hat{A}^2)(1 - 2i\hat{A} - 9\hat{A}^2)}{5 + 7\hat{A}}\right)^\dagger = \frac{(1 + 2i\hat{A}^\dagger - 9\hat{A}^{\dagger 2})(1 - i\hat{A}^\dagger + 3\hat{A}^{\dagger 2})}{5 + 7\hat{A}^\dagger}. \tag{2.86}$$

(c) From (2.81) we immediately infer that the expectation value of a Hermitian operator is real, for it satisfies the following property:

$$\langle\psi \mid \hat{A} \mid \psi\rangle = \langle\psi \mid \hat{A} \mid \psi\rangle^*; \tag{2.87}$$

that is, if $\hat{A}^\dagger = \hat{A}$ then $\langle\psi \mid \hat{A} \mid \psi\rangle$ is real. Similarly, for an anti-Hermitian operator, $\hat{B}^\dagger = -\hat{B}$, we have

$$\langle\psi \mid \hat{B} \mid \psi\rangle = -\langle\psi \mid \hat{B} \mid \psi\rangle^*, \tag{2.88}$$

which means that $\langle\psi \mid \hat{B} \mid \psi\rangle$ is a purely imaginary number.

2.4.3 Projection Operators

An operator $\hat{P}$ is said to be a *projection operator* if it is Hermitian and equal to its own square:

$$\boxed{\text{Projection Operator:} \qquad \hat{P}^\dagger = \hat{P}, \qquad \hat{P}^2 = \hat{P}.} \tag{2.89}$$

The unit operator $\hat{I}$ is a simple example of a projection operator, since $\hat{I}^\dagger = \hat{I}, \quad \hat{I}^2 = \hat{I}$.

Properties of projection operators

- The product of two commuting projection operators, $\hat{P}_1$ and $\hat{P}_2$, is also a projection operator, since

$$(\hat{P}_1\hat{P}_2)^\dagger = \hat{P}_2^\dagger\hat{P}_1^\dagger = \hat{P}_2\hat{P}_1 = \hat{P}_1\hat{P}_2 \text{ and } (\hat{P}_1\hat{P}_2)^2 = \hat{P}_1\hat{P}_2\hat{P}_1\hat{P}_2 = \hat{P}_1^2\hat{P}_2^2 = \hat{P}_1\hat{P}_2. \tag{2.90}$$

- The sum of two projection operators is generally not a projection operator.
- Two projection operators are said to be orthogonal if their product is zero.
- For a sum of projection operators $\hat{P}_1 + \hat{P}_2 + \hat{P}_3 + \cdots$ to be a projection operator, it is necessary and sufficient that these projection operators be mutually orthogonal (i.e., the cross-product terms must vanish).

Example 2.8
Show that the operator $|\psi\rangle\langle\psi|$ is a projection operator only when $|\psi\rangle$ is normalized.

Solution
It is easy to ascertain that the operator $|\psi\rangle\langle\psi|$ is Hermitian, since $(|\psi\rangle\langle\psi|)^\dagger = |\psi\rangle\langle\psi|$. As for the square of this operator, it is given by

$$(|\psi\rangle\langle\psi|)^2 = (|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi|\psi\rangle\langle\psi|. \tag{2.91}$$

Thus, if $|\psi\rangle$ is normalized, we have $(|\psi\rangle\langle\psi|)^2 = |\psi\rangle\langle\psi|$. In sum, if the state $|\psi\rangle$ is normalized, the product of the ket $|\psi\rangle$ with the bra $\langle\psi|$ is a projection operator.

2.4.4 Commutator Algebra

The *commutator* of two operators $\hat{A}$ and $\hat{B}$, denoted by $[\hat{A}, \hat{B}]$, is defined by

$$\boxed{[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A},} \tag{2.92}$$

and the *anticommutator* $\{\hat{A}, \hat{B}\}$ is defined by

$$\boxed{\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}.} \tag{2.93}$$

Two operators are said to commute if their commutator is equal to zero and hence $\hat{A}\hat{B} = \hat{B}\hat{A}$. Any operator commutes with itself:

$$[\hat{A}, \hat{A}] = 0. \tag{2.94}$$

Note that if two operators are Hermitian and their product is also Hermitian, these operators commute:

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger = \hat{B}\hat{A}, \tag{2.95}$$

and since $(\hat{A}\hat{B})^\dagger = \hat{A}\hat{B}$ we have $\hat{A}\hat{B} = \hat{B}\hat{A}$.

As an example, we may mention the commutators involving the x-position operator, $\hat{X}$, and the x-component of the momentum operator, $\hat{P}_x = -i\hbar\partial/\partial x$ (see Equation (2.349) for the proof), as well as the y and the z components

$$\boxed{[\hat{X}, \hat{P}_x] = i\hbar\hat{I}, \qquad [\hat{Y}, \hat{P}_y] = i\hbar\hat{I}, \qquad [\hat{Z}, \hat{P}_z] = i\hbar\hat{I},} \tag{2.96}$$

where $\hat{I}$ is the unit operator.

Properties of commutators
Using the commutator relation (2.92), we can establish the following properties:

- Antisymmetry:
$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}] \tag{2.97}$$
- Linearity:
$$[\hat{A}, \hat{B} + \hat{C} + \hat{D} + \cdots] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}] + [\hat{A}, \hat{D}] + \cdots \tag{2.98}$$
- Hermitian conjugate of a commutator:
$$[\hat{A}, \hat{B}]^\dagger = [\hat{B}^\dagger, \hat{A}^\dagger] \tag{2.99}$$
- Distributivity:
$$[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}] \tag{2.100}$$
$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B} \tag{2.101}$$
- Jacobi identity:
$$[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{B}, [\hat{C}, \hat{A}]] + [\hat{C}, [\hat{A}, \hat{B}]] = 0 \tag{2.102}$$
- By repeated applications of (2.100), we can show that
$$[\hat{A}, \hat{B}^n] = \sum_{j=0}^{n-1} \hat{B}^j [\hat{A}, \hat{B}] \hat{B}^{n-j-1} \tag{2.103}$$
$$[\hat{A}^n, \hat{B}] = \sum_{j=0}^{n-1} \hat{A}^{n-j-1} [\hat{A}, \hat{B}] \hat{A}^j \tag{2.104}$$
- Operators commute with scalars: an operator $\hat{A}$ commutes with any scalar b:
$$[\hat{A}, b] = 0 \tag{2.105}$$

Example 2.9
(a) Show that the commutator of two Hermitian operators is anti-Hermitian.
(b) Evaluate the commutator $[\hat{A}, [\hat{B}, \hat{C}]\hat{D}]$.

Solution

(a) If $\hat{A}$ and $\hat{B}$ are Hermitian, we can write

$$[\hat{A}, \hat{B}]^\dagger = (\hat{A}\hat{B} - \hat{B}\hat{A})^\dagger = \hat{B}^\dagger\hat{A}^\dagger - \hat{A}^\dagger\hat{B}^\dagger = \hat{B}\hat{A} - \hat{A}\hat{B} = -[\hat{A}, \hat{B}]; \tag{2.106}$$

that is, the commutator of $\hat{A}$ and $\hat{B}$ is anti-Hermitian: $[\hat{A}, \hat{B}]^\dagger = -[\hat{A}, \hat{B}]$.

(b) Using the distributivity relation (2.100), we have

$$\begin{aligned} [\hat{A}, [\hat{B}, \hat{C}]\hat{D}] &= [\hat{B}, \hat{C}][\hat{A}, \hat{D}] + [\hat{A}, [\hat{B}, \hat{C}]]\hat{D} \\ &= (\hat{B}\hat{C} - \hat{C}\hat{B})(\hat{A}\hat{D} - \hat{D}\hat{A}) + \hat{A}(\hat{B}\hat{C} - \hat{C}\hat{B})\hat{D} - (\hat{B}\hat{C} - \hat{C}\hat{B})\hat{A}\hat{D} \\ &= \hat{C}\hat{B}\hat{D}\hat{A} - \hat{B}\hat{C}\hat{D}\hat{A} + \hat{A}\hat{B}\hat{C}\hat{D} - \hat{A}\hat{C}\hat{B}\hat{D}. \end{aligned} \tag{2.107}$$

2.4.5 Uncertainty Relation between Two Operators

An interesting application of the commutator algebra is to derive a general relation giving the uncertainties product of two operators, $\hat{A}$ and $\hat{B}$. In particular, we want to give a formal derivation of Heisenberg's uncertainty relations.

Let $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ denote the expectation values of two Hermitian operators $\hat{A}$ and $\hat{B}$ with respect to a normalized state vector $|\psi\rangle$: $\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle$ and $\langle\hat{B}\rangle = \langle\psi|\hat{B}|\psi\rangle$. Introducing the operators $\Delta\hat{A}$ and $\Delta\hat{B}$,

$$\Delta\hat{A} = \hat{A} - \langle\hat{A}\rangle, \qquad \Delta\hat{B} = \hat{B} - \langle\hat{B}\rangle, \tag{2.108}$$

we have $(\Delta\hat{A})^2 = \hat{A}^2 - 2\hat{A}\langle\hat{A}\rangle + \langle\hat{A}\rangle^2$ and $(\Delta\hat{B})^2 = \hat{B}^2 - 2\hat{B}\langle\hat{B}\rangle + \langle\hat{B}\rangle^2$, and hence

$$\langle\psi|(\Delta\hat{A})^2|\psi\rangle = \langle(\Delta\hat{A})^2\rangle = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2, \qquad \langle(\Delta\hat{B})^2\rangle = \langle\hat{B}^2\rangle - \langle\hat{B}\rangle^2, \tag{2.109}$$

where $\langle\hat{A}^2\rangle = \langle\psi|\hat{A}^2|\psi\rangle$ and $\langle\hat{B}^2\rangle = \langle\psi|\hat{B}^2|\psi\rangle$. The *uncertainties* ΔA and ΔB are defined by

$$\boxed{\Delta A = \sqrt{\langle(\Delta\hat{A})^2\rangle} = \sqrt{\langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2}, \qquad \Delta B = \sqrt{\langle(\Delta\hat{B})^2\rangle} = \sqrt{\langle\hat{B}^2\rangle - \langle\hat{B}\rangle^2}.} \tag{2.110}$$

Let us write the action of the operators (2.108) on any state $|\psi\rangle$ as follows:

$$|\chi\rangle = \Delta\hat{A}|\psi\rangle = \left(\hat{A} - \langle\hat{A}\rangle\right)|\psi\rangle, \qquad |\phi\rangle = \Delta\hat{B}|\psi\rangle = \left(\hat{B} - \langle\hat{B}\rangle\right)|\psi\rangle. \tag{2.111}$$

The Schwarz inequality for the states $|\chi\rangle$ and $|\phi\rangle$ is given by

$$\langle\chi|\chi\rangle\langle\phi|\phi\rangle \geq \left|\langle\chi|\phi\rangle\right|^2. \tag{2.112}$$

Since $\hat{A}$ and $\hat{B}$ are Hermitian, $\Delta\hat{A}$ and $\Delta\hat{B}$ must also be Hermitian: $\Delta\hat{A}^\dagger = \hat{A}^\dagger - \langle\hat{A}\rangle = \hat{A} - \langle\hat{A}\rangle = \Delta\hat{A}$ and $\Delta\hat{B}^\dagger = \hat{B} - \langle\hat{B}\rangle = \Delta\hat{B}$. Thus, we can show the following three relations:

$$\langle\chi|\chi\rangle = \langle\psi|(\Delta\hat{A})^2|\psi\rangle, \quad \langle\phi|\phi\rangle = \langle\psi|(\Delta\hat{B})^2|\psi\rangle, \quad \langle\chi|\phi\rangle = \langle\psi|\Delta\hat{A}\Delta\hat{B}|\psi\rangle. \tag{2.113}$$

For instance, since $\Delta\hat{A}^\dagger = \Delta\hat{A}$ we have $\langle\chi \mid \chi\rangle = \langle\psi \mid \Delta\hat{A}^\dagger\Delta\hat{A} \mid \psi\rangle = \langle\psi \mid (\Delta\hat{A})^2 \mid \psi\rangle = \langle(\Delta\hat{A})^2\rangle$. Hence, the Schwarz inequality (2.112) becomes

$$\langle(\Delta\hat{A})^2\rangle\langle(\Delta\hat{B})^2\rangle \geq \left|\langle\Delta\hat{A}\Delta\hat{B}\rangle\right|^2. \tag{2.114}$$

Notice that the last term $\Delta\hat{A}\Delta\hat{B}$ of this equation can be written as

$$\Delta\hat{A}\Delta\hat{B} = \frac{1}{2}[\Delta\hat{A},\ \Delta\hat{B}] + \frac{1}{2}\{\Delta\hat{A},\ \Delta\hat{B}\} = \frac{1}{2}[\hat{A},\ \hat{B}] + \frac{1}{2}\{\Delta\hat{A},\ \Delta\hat{B}\}, \tag{2.115}$$

where we have used the fact that $[\Delta\hat{A},\ \Delta\hat{B}] = [\hat{A},\ \hat{B}]$. Since $[\hat{A},\ \hat{B}]$ is anti-Hermitian and $\{\Delta\hat{A},\ \Delta\hat{B}\}$ is Hermitian and since the expectation value of a Hermitian operator is real and that the expectation value of an anti-Hermitian operator is imaginary (see Example 2.7), the expectation value $\langle\Delta\hat{A}\Delta\hat{B}\rangle$ of (2.115) becomes equal to the sum of a real part $\langle\{\Delta\hat{A},\ \Delta\hat{B}\}\rangle/2$ and an imaginary part $\langle[\hat{A},\ \hat{B}]\rangle/2$; hence

$$\left|\langle\Delta\hat{A}\Delta\hat{B}\rangle\right|^2 = \frac{1}{4}\left|\langle[\hat{A},\ \hat{B}]\rangle\right|^2 + \frac{1}{4}\left|\langle\{\Delta\hat{A},\ \Delta\hat{B}\}\rangle\right|^2. \tag{2.116}$$

Since the last term is a positive real number, we can infer the following relation:

$$\left|\langle\Delta\hat{A}\Delta\hat{B}\rangle\right|^2 \geq \frac{1}{4}\left|\langle[\hat{A},\ \hat{B}]\rangle\right|^2. \tag{2.117}$$

Comparing equations (2.114) and (2.117), we conclude that

$$\langle(\Delta\hat{A})^2\rangle\langle(\Delta\hat{B})^2\rangle \geq \frac{1}{4}\left|\langle[\hat{A},\ \hat{B}]\rangle\right|^2, \tag{2.118}$$

which (by taking its square root) can be reduced to

$$\boxed{\Delta A\Delta B \geq \frac{1}{2}\left|\langle[\hat{A},\ \hat{B}]\rangle\right|.} \tag{2.119}$$

This uncertainty relation plays an important role in the formalism of quantum mechanics. Its application to position and momentum operators leads to the Heisenberg uncertainty relations, which represent one of the cornerstones of quantum mechanics; see the next example.

Example 2.10 (Heisenberg uncertainty relations)
Find the uncertainty relations between the components of the position and the momentum operators.

Solution
By applying (2.119) to the x-components of the position operator $\hat{X}$, and the momentum operator $\hat{P}_x$, we obtain $\Delta x\Delta p_x \geq \frac{1}{2}\,|\,\langle[\hat{X},\ \hat{P}_x]\rangle\,|$. But since $[\hat{X},\ \hat{P}_x] = i\hbar\hat{I}$, we have $\Delta x\Delta p_x \geq \hbar/2$; the uncertainty relations for the $y-$ and $z-$ components follow immediately:

$$\boxed{\Delta x\Delta p_x \geq \frac{\hbar}{2}, \qquad \Delta y\Delta p_y \geq \frac{\hbar}{2}, \qquad \Delta z\Delta p_z \geq \frac{\hbar}{2}.} \tag{2.120}$$

These are the Heisenberg uncertainty relations.

2.4.6 Functions of Operators

Let $F(\hat{A})$ be a function of an operator $\hat{A}$. If $\hat{A}$ is a linear operator, we can Taylor expand $F(\hat{A})$ in a power series of $\hat{A}$:

$$F(\hat{A}) = \sum_{n=0}^{\infty} a_n \hat{A}^n, \tag{2.121}$$

where a_n is just an expansion coefficient. As an illustration of an operator function, consider $e^{a\hat{A}}$, where a is a scalar which can be complex or real. We can expand it as follows:

$$e^{a\hat{A}} = \sum_{n=0}^{\infty} \frac{a^n}{n!}\hat{A}^n = \hat{I} + a\hat{A} + \frac{a^2}{2!}\hat{A}^2 + \frac{a^3}{3!}\hat{A}^3 + \cdots. \tag{2.122}$$

Commutators involving function operators
If $\hat{A}$ commutes with another operator $\hat{B}$, then $\hat{B}$ commutes with any operator function that depends on $\hat{A}$:

$$[\hat{A}, \hat{B}] = 0 \implies [\hat{B}, F(\hat{A})] = 0; \tag{2.123}$$

in particular, $F(\hat{A})$ commutes with $\hat{A}$ and with any other function, $G(\hat{A})$, of $\hat{A}$:

$$[\hat{A}, F(\hat{A})] = 0, \qquad [\hat{A}^n, F(\hat{A})] = 0, \qquad [F(\hat{A}), G(\hat{A})] = 0. \tag{2.124}$$

Hermitian adjoint of function operators
The adjoint of $F(\hat{A})$ is given by

$$[F(\hat{A})]^\dagger = F^*(\hat{A}^\dagger). \tag{2.125}$$

Note that if $\hat{A}$ is Hermitian, $F(\hat{A})$ is not necessarily Hermitian; $F(\hat{A})$ will be Hermitian only if F is a real function and $\hat{A}$ is Hermitian. An example is

$$(e^{\hat{A}})^\dagger = e^{\hat{A}^\dagger}, \qquad (e^{i\hat{A}})^\dagger = e^{-i\hat{A}^\dagger}, \qquad (e^{ia\hat{A}})^\dagger = e^{-ia^*\hat{A}^\dagger}, \tag{2.126}$$

where a is a complex number. So if $\hat{A}$ is Hermitian, an operator function which can be expanded as $F(\hat{A}) = \sum_{n=0}^{\infty} a_n \hat{A}^n$ will be Hermitian only if the expansion coefficients a_n are real numbers. But in general, $F(\hat{A})$ is not Hermitian even if $\hat{A}$ is Hermitian, since

$$F^*(\hat{A}^\dagger) = \sum_{n=0}^{\infty} a_n^* (\hat{A}^\dagger)^n. \tag{2.127}$$

Relations involving function operators
Note that

$$[\hat{A}, \hat{B}] \neq 0 \implies [\hat{B}, F(\hat{A})] \neq 0; \tag{2.128}$$

in particular, $e^{\hat{A}}e^{\hat{B}} \neq e^{\hat{A}+\hat{B}}$. Using (2.122) we can ascertain that

$$e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}}e^{[\hat{A}, \hat{B}]/2}, \tag{2.129}$$

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A}, \hat{B}] + \frac{1}{2!}[\hat{A}, [\hat{A}, \hat{B}]] + \frac{1}{3!}[\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]] + \cdots. \tag{2.130}$$

2.4.7 Inverse of an Operator and Unitary Operators

Inverse of an operator: Assuming it exists[3] the *inverse* $\hat{A}^{-1}$ of a linear operator $\hat{A}$ is defined by the relation

$$\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{I}, \tag{2.131}$$

where $\hat{I}$ is the unit operator, the operator that leaves any state $|\psi\rangle$ unchanged.

Quotient of two operators: Dividing an operator $\hat{A}$ by another operator $\hat{B}$ (provided that the inverse $\hat{B}^{-1}$ exists) is equivalent to multiplying $\hat{A}$ by $\hat{B}^{-1}$:

$$\frac{\hat{A}}{\hat{B}} = \hat{A}\hat{B}^{-1}. \tag{2.132}$$

The side on which the quotient is taken matters:

$$\frac{\hat{A}}{\hat{B}} = \hat{A}\frac{\hat{I}}{\hat{B}} = \hat{A}\hat{B}^{-1} \quad \text{and} \quad \frac{\hat{I}}{\hat{B}}\hat{A} = \hat{B}^{-1}\hat{A}. \tag{2.133}$$

In general, we have $\hat{A}\hat{B}^{-1} \neq \hat{B}^{-1}\hat{A}$. For an illustration of these ideas, see Problem 2.13. We may mention here the following properties about the inverse of operators:

$$\left(\hat{A}\hat{B}\hat{C}\hat{D}\right)^{-1} = \hat{D}^{-1}\hat{C}^{-1}\hat{B}^{-1}\hat{A}^{-1}, \qquad \left(\hat{A}^{n}\right)^{-1} = \left(\hat{A}^{-1}\right)^{n}. \tag{2.134}$$

Unitary operators: A linear operator $\hat{U}$ is said to be *unitary* if its inverse $\hat{U}^{-1}$ is equal to its adjoint $\hat{U}^{\dagger}$:

$$\boxed{\hat{U} \text{ is a unitary operator:} \qquad \hat{U}^{\dagger} = \hat{U}^{-1} \qquad \text{or} \qquad \hat{U}\hat{U}^{\dagger} = \hat{U}^{\dagger}\hat{U} = \hat{I}.} \tag{2.135}$$

The product of two unitary operators is also unitary, since

$$(\hat{U}\hat{V})(\hat{U}\hat{V})^{\dagger} = (\hat{U}\hat{V})(\hat{V}^{\dagger}\hat{U}^{\dagger}) = \hat{U}(\hat{V}\hat{V}^{\dagger})\hat{U}^{\dagger} = \hat{U}\hat{U}^{\dagger} = \hat{I}, \tag{2.136}$$

or $(\hat{U}\hat{V})^{\dagger} = (\hat{U}\hat{V})^{-1}$. This result can be generalized to any number of operators; the product of a number of unitary operators is also unitary, since

$$\begin{aligned}(\hat{A}\hat{B}\hat{C}\hat{D}\cdots)(\hat{A}\hat{B}\hat{C}\hat{D}\cdots)^{\dagger} &= \hat{A}\hat{B}\hat{C}\hat{D}(\cdots)\hat{D}^{\dagger}\hat{C}^{\dagger}\hat{B}^{\dagger}\hat{A}^{\dagger} = \hat{A}\hat{B}\hat{C}(\hat{D}\hat{D}^{\dagger})\hat{C}^{\dagger}\hat{B}^{\dagger}\hat{A}^{\dagger} \\ &= \hat{A}\hat{B}(\hat{C}\hat{C}^{\dagger})\hat{B}^{\dagger}\hat{A}^{\dagger} = \hat{A}(\hat{B}\hat{B}^{\dagger})\hat{A}^{\dagger} \\ &= \hat{A}\hat{A}^{\dagger} = \hat{I},\end{aligned} \tag{2.137}$$

or $(\hat{A}\hat{B}\hat{C}\hat{D}\cdots)^{\dagger} = (\hat{A}\hat{B}\hat{C}\hat{D}\cdots)^{-1}$.

---

[3]Not every operator has an inverse, just as in the case of matrices. The inverse of a matrix exists only when its determinant is nonzero.

Example 2.11 (Unitary operator)
What conditions must the parameter ϵ and the operator $\hat{G}$ satisfy so that the operator $\hat{U} = e^{i\epsilon\hat{G}}$ is unitary?

Solution
Clearly, if ϵ is real and $\hat{G}$ is Hermitian, the operator $e^{i\epsilon\hat{G}}$ would be unitary. Using the property $[F(\hat{A})]^\dagger = F^*(\hat{A}^\dagger)$, we see that

$$(e^{i\epsilon\hat{G}})^\dagger = e^{-i\epsilon\hat{G}} = (e^{i\epsilon\hat{G}})^{-1}, \tag{2.138}$$

that is, $\hat{U}^\dagger = \hat{U}^{-1}$.

2.4.8 Antilinear and Antiunitary Operators, Wigner Theorem

In this Subsection, we are going to look at antiunitary operators as well as at the Wigner theorem that deals with their properties. These topis are important in quantum mechanics, especially when we study the *time reversal symmetry* in Chapter 12.

Antilinear Operators:
An operator $\hat{A}$ is said to be *antilinear* or, equivalently, *conjugate linear* if it obeys the distributive law and, like all operators, it commutes with constants, but it takes the complex conjugate of the numbers on its right-hand side. That is, an operator $\hat{A}$ is antilinear if, for any states $|\psi_1\rangle$ and $|\psi_2\rangle$ and any complex numbers a_1 and a_2, we have

$$\boxed{\hat{A}(a_1|\psi_1\rangle + a_2|\psi_2\rangle) = a_1^*\hat{A}|\psi_1\rangle + a_2^*\hat{A}|\psi_2\rangle.} \tag{2.139}$$

In particular, if $a_2 = 0$, we can write this commutation relation: $\hat{A}a_1 = a_1^*\hat{A}$, which can be rewritten in more general terms as

$$\boxed{\text{For an antilinear operator } \hat{A} \text{ and a complex number } c: \qquad \hat{A}c = c^*\hat{A}.} \tag{2.140}$$

This leads to $\hat{A}i = -i\hat{A}$ and $\hat{A}b = b\hat{A}$ where b is a real number. It is important to mention that *an antilinear operator acts from the left to the right only.*

We should note that the product of N antilinear operators is linear for even N and antilinear for odd N. In particular, if $\hat{A}|\psi\rangle = a|\psi\rangle$, where a is a complex number, then $\hat{A}^2$ is linear, since

$$\hat{A}^2|\psi\rangle = \hat{A}(a|\psi\rangle) = aa^*|\psi\rangle. \tag{2.141}$$

Hence, if $|\psi\rangle$ is an eigenvector of $\hat{A}$ with eigenvalue a, the corresponding eigenvalue of $\hat{A}^2$ is real and positive, since $aa^* \equiv |a|^2 \geq 0$.

Antiunitary Operators:
Let us consider an operator $\hat{T}$ which transforms two states $|\psi_1\rangle$ and $|\psi_2\rangle$ into $|\psi_1'\rangle$ and $|\psi_2'\rangle$, respectively:

$$\hat{T}|\psi_1\rangle = |\psi_1'\rangle, \qquad \hat{T}|\psi_2\rangle = |\psi_2'\rangle. \tag{2.142}$$

The operator $\hat{T}$ is said to be *antiunitary* if, for any states $|\psi_1\rangle$ and $|\psi_2\rangle$ and any complex numbers a_1 and a_2, it satisfies the following two conditions

$$\hat{T}(a_1|\psi_1\rangle + a_2|\psi_2\rangle) = a_1^*|\psi_1'\rangle + a_2^*|\psi_2'\rangle \tag{2.143}$$

$$\langle\psi_1'|\psi_2'\rangle = \langle\hat{T}\psi_1|\hat{T}\psi_2\rangle = \langle\psi_1|\psi_2\rangle^* = \langle\psi_2|\psi_1\rangle. \tag{2.144}$$

So, an antilinear operator that satisfies the condition $\langle\hat{T}\psi_1|\hat{T}\psi_2\rangle = \langle\psi_2|\psi_1\rangle$ is called antiunitary. The requirement of invariance of an operator $\hat{H}$ under $\hat{T}$ is expressed as follows

$$\hat{T}\hat{H}\hat{T}^\dagger = \hat{H}. \tag{2.145}$$

We can also show that *the product of two antiunitary operators is a unitary operator* (see Example 2.12).

Properties of antiunitary operators:

- *The product of an antiunitary operator times a unitary operator is an antiunitary operator;* for a proof see Problem 2.15 on Page 2.15.
- *The product of a unitary operator times a complex conjugation operator is an antiunitary operator.* The proof is as follows. First, since the complex conjugation operator $\hat{K}$ only acts on the complex numbers on its right-hand side, we have: $\hat{K}a = a^*\hat{K}$. The operator $\hat{T} = \hat{U}\hat{K}$ is clearly antilinear:

$$\begin{aligned}\hat{T}(a_1|\psi_1\rangle + a_2|\psi_2\rangle) &= \hat{U}\hat{K}(a_1|\psi_1\rangle) + \hat{U}\hat{K}(a_2|\psi_2\rangle) = \hat{U}a_1^*(\hat{K}|\psi_1\rangle) + \hat{U}a_2^*(\hat{K}|\psi_2\rangle)\\ &= a_1^*(\hat{U}\hat{K}|\psi_1\rangle) + a_2^*(\hat{U}\hat{K}|\psi_2\rangle)\\ &= a_1^*\hat{T}|\psi_1\rangle + a_2^*\hat{T}|\psi_2\rangle\\ &= a_1^*|\psi_1'\rangle + a_2^*|\psi_2'\rangle,\end{aligned} \tag{2.146}$$

  where we have used the relation $\hat{U}a_j^* = a_j^*\hat{U}$ $(j = 1, 2)$ in the second line, since unitary operators commute with complex numbers. Second, we can verify that the operator $\hat{T} = \hat{U}\hat{K}$ is antiunitary:

$$\begin{aligned}\langle\psi_1'|\psi_2'\rangle &= \big(\langle\psi_1|\hat{T}^\dagger\big)\big(\hat{T}|\psi_2\rangle\big) =\\ &= \big(\langle\psi_1|\hat{U}^\dagger\big)\hat{K}^\dagger\hat{K}\big(\hat{U}|\psi_2\rangle\big)\\ &= \big(\langle\psi_1|\hat{U}^\dagger\hat{U}|\psi_2\rangle\big)^*\\ &= \langle\psi_1|\psi_2\rangle^*\\ &= \langle\psi_2|\psi_1\rangle,\end{aligned} \tag{2.147}$$

  where we have used the relation $\hat{U}^\dagger\hat{U} = I$ in the fourth line, since $\hat{U}$ is a unitary operator. Equations (2.146) and (2.147) show that the operator $\hat{T} = \hat{U}\hat{K}$ is both antilinear and antiunitary.

- Applying the complex conjugation operator twice in a succession will give back the original expression:

$$\hat{K}^2 = I \quad \Longrightarrow \quad \hat{K} = \hat{K}^{-1} = \hat{K}, \tag{2.148}$$

where I is the identity operator. For instance, applying $\hat{K}$ twice on a complex number c would yield: $\hat{K}^2 c = \hat{K} c^* = c$.

- The complex conjugation operator $\hat{K}$ is antiunitray, since

$$\langle \hat{K}\psi_1 \mid \hat{K}\psi_2 \rangle = \langle \psi_1 \mid \psi_2 \rangle^* = \langle \psi_2 \mid \psi_1 \rangle. \tag{2.149}$$

- The Hermitian conjugate (or adjoint) of an antiunitary operator is also antiunitary:

$$\hat{T}^\dagger \hat{T} = \hat{T}\hat{T}^\dagger = \boldsymbol{I}. \tag{2.150}$$

This relation should not be confused with the property of a unitary operator $\hat{U}$: $\hat{U}^\dagger \hat{U} = \hat{U}\hat{U}^\dagger = I$; the antiunitary operator $\hat{T}$ is antilinear, while the unitary operator $\hat{U}$ is linear.

Eugene Wigner's Theorem:
According to Wigner's theorem[4], *any symmetry transformation of Hilbert space is represented by either a linear unitary or an antilinear antiunitary operator.*

As a result, and considering a symmetry S that transforms two states $\mid \psi_1 \rangle$ and $\mid \psi_2 \rangle$ into $\mid \psi_1' \rangle$ and $\mid \psi_2' \rangle$, respectively

$$\mid \psi_1 \rangle \xrightarrow{S} \mid \psi_1' \rangle \qquad \text{and} \qquad \mid \psi_2 \rangle \xrightarrow{S} \mid \psi_2' \rangle \tag{2.151}$$

we expect the absolute value of the scalar product between these two states to be invariant under the transformation S:

$$\left| \langle \psi_1' \mid \psi_2' \rangle \right| = \left| \langle \psi_1 \mid \psi_2 \rangle \right|. \tag{2.152}$$

This invariance condition can be achieved in two ways:

- by means of a symmetry that is represented by a *unitary* operator $\hat{U}$

$$\langle \psi_1 \mid \psi_2 \rangle \xrightarrow{\hat{U}} \langle \psi_1' \mid \psi_2' \rangle = \langle \psi_1 \mid \hat{U}^\dagger \hat{U} \mid \psi_2 \rangle = \langle \psi_1 \mid \psi_2 \rangle; \tag{2.153}$$

that is, $\hat{U}^\dagger \hat{U} = \hat{I}$ where $\hat{I}$ is the unit operator.

- or by means of a symmetry that is represented by an *antiunitary* operator $\hat{T}$

$$\langle \psi_1 \mid \psi_2 \rangle \xrightarrow{\hat{T}} \langle \psi_1' \mid \psi_2' \rangle = \langle \psi_1 \mid \psi_2 \rangle^* = \langle \psi_2 \mid \psi_1 \rangle. \tag{2.154}$$

The two conditions of antilinearity and antiunitarity that must be satisfied by this transformation are listed in Eqs. (2.143)–(2.144).

From (2.153) and (2.154), we see that both unitary and antiunitary transformations preserve probabilities, since

$$\left| \langle \psi_1' \mid \psi_2' \rangle \right|^2 = \left| \langle \psi_1 \mid \psi_2 \rangle \right|^2. \tag{2.155}$$

[4]The Wigner theorem first appeared in a 1931 book written by E. Wigner (in German): *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*, Vieweg, Braunscheig, Germany (1931). English translation: E. Wigner, *Group theory and its applications to quantum mechanics of atomic spectra*, Academic Press Inc., New York (1959).

One of the consequences of Wigner's theorem is the fact that *every inversion operator which preserves the transition probabilities is either unitary or antiunitary.* Additionally, *any unitary operator is linear and any antiunitary operator is antilinear.* Also, and as shown above, *every antiunitary operator may be decomposed into the product of a unitary operator times a complex conjugation operator.* This result will be of particular usefulness in Chapter 12 when we deal with the *time reversal operator.*

We should note that in non-relativistic quantum mechanics, which is treated in the first eleven (11) chapters of this text, we deal primarily with unitary operators; these are suitable for describing observables with positive eigenvalues (e.g., energies). However, when we deal with relativistic quantum mechanics in Chapter 12, we will describe the time reversal symmetry in terms of antiunitary operators so as to avoid encountering negative energies. So, since antilinear and antiunitary transformations are related to time inversion, we will see them again in Chapter 12.

Example 2.12 (The product of two antiunitary operators is a unitary operator)
Show that the product of two antiunitary operators is a unitary operator.

Solution

Consider an operator $\hat{K} = \hat{T}\hat{F}$ where $\hat{T}$ and $\hat{F}$ are two antiunitary operators which transform two states ϕ_1 and ϕ_2 into ϕ_1' and ϕ_2', respectively:

$$\hat{K} \mid \phi_1\rangle = \mid \phi_1'\rangle, \qquad \hat{K} \mid \phi_2\rangle = \mid \phi_2'\rangle. \tag{2.156}$$

For any two states $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ and any two complex numbers a_1 and a_2, we can write

$$\begin{aligned} \hat{K}\left(a_1 \mid \phi_1\rangle + a_2 \mid \phi_2\rangle\right) &= \hat{T}\hat{F}(a_1 \mid \phi_1\rangle) + \hat{T}\hat{F}(a_2 \mid \phi_2\rangle) = \hat{T}(a_1^*\hat{F} \mid \phi_1\rangle) + \hat{T}(a_2^*\hat{F} \mid \phi_2\rangle) \\ &= a_1\hat{T}\hat{F} \mid \phi_1\rangle + a_2\hat{T}\hat{F} \mid \phi_2\rangle \\ &= a_1(\hat{K} \mid \phi_1\rangle) + a_2(\hat{K} \mid \phi_2\rangle) \\ &= a_1 \mid \phi_1'\rangle + a_2 \mid \phi_2'\rangle. \end{aligned} \tag{2.157}$$

Additionally, we have

$$\begin{aligned} \langle \phi_1' \mid \phi_2'\rangle &= \left(\langle \phi_1 \mid \hat{K}^\dagger\right)\left(\hat{K} \mid \phi_2\rangle\right) = \\ &= \left(\langle \phi_1 \mid \hat{F}^\dagger\right)\hat{T}^\dagger\hat{T}\left(\hat{F} \mid \phi_2\rangle\right) \\ &= \left(\langle \phi_1 \mid \hat{F}^\dagger\hat{F} \mid \phi_2\rangle\right)^* \\ &= \left(\langle \phi_1 \mid \phi_2\rangle^*\right)^* \\ &= \langle \phi_1 \mid \phi_2\rangle. \end{aligned} \tag{2.158}$$

Relation (2.157) shows that $\hat{K}$ is linear and (2.158) implies that $\hat{K}^\dagger\hat{K} = I_4$. Consequently, the operator $\hat{K}$ is unitary. That is, the product of two antiunitary operators is a unitary operator

---

2.4.9 Eigenvalues and Eigenvectors of an Operator

Having studied the properties of operators and states, we are now ready to discuss how to find the eigenvalues and eigenvectors of an operator.

A state vector $|\psi\rangle$ is said to be an *eigenvector* (also called an eigenket or eigenstate) of an operator $\hat{A}$ if the application of $\hat{A}$ to $|\psi\rangle$ gives

$$\hat{A}|\psi\rangle = a|\psi\rangle, \tag{2.159}$$

where a is a complex number, called an *eigenvalue* of $\hat{A}$. This equation is known as the *eigenvalue equation*, or *eigenvalue problem*, of the operator $\hat{A}$. Its solutions yield the eigenvalues and eigenvectors of $\hat{A}$. In Section 2.5.3 we will see how to solve the eigenvalue problem in a discrete basis.

A simple example is the eigenvalue problem for the unity operator $\hat{I}$:

$$\hat{I}|\psi\rangle = |\psi\rangle. \tag{2.160}$$

This means that all vectors are eigenvectors of $\hat{I}$ with one eigenvalue, 1. Note that

$$\hat{A}|\psi\rangle = a|\psi\rangle \Longrightarrow \hat{A}^n|\psi\rangle = a^n|\psi\rangle \quad \text{and} \quad F(\hat{A})|\psi\rangle = F(a)|\psi\rangle. \tag{2.161}$$

For instance, we have

$$\hat{A}|\psi\rangle = a|\psi\rangle \quad \Longrightarrow \quad e^{i\hat{A}}|\psi\rangle = e^{ia}|\psi\rangle. \tag{2.162}$$

Example 2.13 (Eigenvalues of the inverse of an operator)
Show that if $\hat{A}^{-1}$ exists, the eigenvalues of $\hat{A}^{-1}$ are just the inverses of those of $\hat{A}$.

Solution
Since $\hat{A}^{-1}\hat{A} = \hat{I}$ we have on the one hand

$$\hat{A}^{-1}\hat{A}|\psi\rangle = |\psi\rangle, \tag{2.163}$$

and on the other hand

$$\hat{A}^{-1}\hat{A}|\psi\rangle = \hat{A}^{-1}(\hat{A}|\psi\rangle) = a\hat{A}^{-1}|\psi\rangle. \tag{2.164}$$

Combining the previous two equations, we obtain

$$a\hat{A}^{-1}|\psi\rangle = |\psi\rangle, \tag{2.165}$$

hence

$$\hat{A}^{-1}|\psi\rangle = \frac{1}{a}|\psi\rangle. \tag{2.166}$$

This means that $|\psi\rangle$ is also an eigenvector of $\hat{A}^{-1}$ with eigenvalue $1/a$. That is, if $\hat{A}^{-1}$ exists, then

$$\boxed{\hat{A}|\psi\rangle = a|\psi\rangle \quad \Longrightarrow \quad \hat{A}^{-1}|\psi\rangle = \frac{1}{a}|\psi\rangle.} \tag{2.167}$$

Some useful theorems pertaining to the eigenvalue problem

Theorem 2.1 *For a Hermitian operator, all of its eigenvalues are real and the eigenvectors corresponding to different eigenvalues are orthogonal.*

$$\text{If } \hat{A}^\dagger = \hat{A}, \quad \hat{A} \mid \phi_n\rangle = a_n \mid \phi_n\rangle \implies a_n = \text{real number, and } \langle \phi_m \mid \phi_n\rangle = \delta_{mn}. \tag{2.168}$$

Proof of Theorem 2.1
Note that

$$\hat{A} \mid \phi_n\rangle = a_n \mid \phi_n\rangle \implies \langle \phi_m \mid \hat{A} \mid \phi_n\rangle = a_n \langle \phi_m \mid \phi_n\rangle, \tag{2.169}$$

and

$$\langle \phi_m \mid \hat{A}^\dagger = a_m^* \langle \phi_m \mid \implies \langle \phi_m \mid \hat{A}^\dagger \mid \phi_n\rangle = a_m^* \langle \phi_m \mid \phi_n\rangle. \tag{2.170}$$

Subtracting (2.170) from (2.169) and using the fact that $\hat{A}$ is Hermitian, $\hat{A} = \hat{A}^\dagger$, we have

$$(a_n - a_m^*)\langle \phi_m \mid \phi_n\rangle = 0. \tag{2.171}$$

Two cases must be considered separately:

- Case $m = n$: since $\langle \phi_n \mid \phi_n\rangle > 0$, we must have $a_n = a_n^*$; hence the eigenvalues a_n must be real.
- Case $m \neq n$: since in general $a_n \neq a_m^*$, we must have $\langle \phi_m \mid \phi_n\rangle = 0$; that is, $\mid \phi_m\rangle$ and $\mid \phi_n\rangle$ must be orthogonal.

Theorem 2.2 *The eigenstates of a Hermitian operator define a complete set of mutually orthonormal basis states. The operator is diagonal in this eigenbasis with its diagonal elements equal to the eigenvalues. This basis set is unique if the operator has no degenerate eigenvalues and not unique (in fact it is infinite) if there is any degeneracy.*

Theorem 2.3 *If two Hermitian operators, $\hat{A}$ and $\hat{B}$, commute and if $\hat{A}$ has no degenerate eigenvalue, then each eigenvector of $\hat{A}$ is also an eigenvector of $\hat{B}$. In addition, we can construct a common orthonormal basis that is made of the joint eigenvectors of $\hat{A}$ and $\hat{B}$.*

Proof of Theorem 2.3
Since $\hat{A}$ is Hermitian with no degenerate eigenvalue, to each eigenvalue of $\hat{A}$ there corresponds only one eigenvector. Consider the equation

$$\hat{A} \mid \phi_n\rangle = a_n \mid \phi_n\rangle. \tag{2.172}$$

Since $\hat{A}$ commutes with $\hat{B}$ we can write

$$\hat{B}\hat{A} \mid \phi_n\rangle = \hat{A}\hat{B} \mid \phi_n\rangle \quad \text{or} \quad \hat{A}(\hat{B} \mid \phi_n)) = a_n(\hat{B} \mid \phi_n)); \tag{2.173}$$

that is, $(\hat{B} \mid \phi_n))$ is an eigenvector of $\hat{A}$ with eigenvalue a_n. But since this eigenvector is unique (apart from an arbitrary phase constant), the ket $\mid \phi_n\rangle$ must also be an eigenvector of $\hat{B}$:

$$\hat{B} \mid \phi_n\rangle = b_n \mid \phi_n\rangle. \tag{2.174}$$

Since each eigenvector of $\hat{A}$ is also an eigenvector of $\hat{B}$ (and vice versa), both of these operators must have a common basis. This basis is unique; it is made of the joint eigenvectors of $\hat{A}$ and $\hat{B}$. This theorem also holds for any number of mutually commuting Hermitian operators.

Now, if a_n is a degenerate eigenvalue, we can only say that $\hat{B} \mid \phi_n\rangle$ is an eigenvector of $\hat{A}$ with eigenvalue a_n; $\mid \phi_n\rangle$ is not necessarily an eigenvector of $\hat{B}$. If one of the operators is degenerate, there exist an infinite number of orthonormal basis sets that are common to these two operators; that is, the joint basis does exist and it is not unique.

Theorem 2.4 *The eigenvalues of an anti-Hermitian operator are either purely imaginary or equal to zero.*

Theorem 2.5 *The eigenvalues of a unitary operator are complex numbers of moduli equal to one; the eigenvectors of a unitary operator that has no degenerate eigenvalues are mutually orthogonal.*

Proof of Theorem 2.5
Let $\mid \phi_n\rangle$ and $\mid \phi_m\rangle$ be eigenvectors to the unitary operator $\hat{U}$ with eigenvalues a_n and a_m, respectively. We can write

$$(\langle \phi_m \mid \hat{U}^\dagger)(\hat{U} \mid \phi_n\rangle) = a_m^* a_n \langle \phi_m \mid \phi_n\rangle. \tag{2.175}$$

Since $\hat{U}^\dagger \hat{U} = \hat{I}$ this equation can be rewritten as

$$(a_m^* a_n - 1)\langle \phi_m \mid \phi_n\rangle = 0, \tag{2.176}$$

which in turn leads to the following two cases:

- Case $n = m$: since $\langle \phi_n \mid \phi_n\rangle > 0$ then $a_n^* a_n = \mid a_n \mid^2 = 1$, and hence $\mid a_n \mid = 1$.
- Case $n \neq m$: the only possibility for this case is that $\mid \phi_m\rangle$ and $\mid \phi_n\rangle$ are orthogonal, $\langle \phi_m \mid \phi_n\rangle = 0$.

2.4.10 Infinitesimal and Finite Unitary Transformations

We want to study here how quantities such as kets, bras, operators, and scalars transform under unitary transformations. A unitary transformation is the application of a unitary operator $\hat{U}$ to one of these quantities.

2.4.10.1 Unitary Transformations

Kets $\mid \psi\rangle$ and bras $\langle \psi \mid$ transform as follows:

$$\mid \psi'\rangle = \hat{U} \mid \psi\rangle, \qquad \langle \psi' \mid = \langle \psi \mid \hat{U}^\dagger. \tag{2.177}$$

Let us now find out how operators transform under unitary transformations. Since the transform of $\hat{A} \mid \psi\rangle = \mid \phi\rangle$ is $\hat{A}' \mid \psi'\rangle = \mid \phi'\rangle$, we can rewrite $\hat{A}' \mid \psi'\rangle = \mid \phi'\rangle$ as $\hat{A}'\hat{U} \mid \psi\rangle = \hat{U} \mid \phi\rangle = \hat{U}\hat{A} \mid \psi\rangle$ which, in turn, leads to $\hat{A}'\hat{U} = \hat{U}\hat{A}$. Multiplying both sides of $\hat{A}'\hat{U} = \hat{U}\hat{A}$ by $\hat{U}^\dagger$ and since $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}$, we have

$$\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger, \qquad \hat{A} = \hat{U}^\dagger \hat{A}' \hat{U}. \tag{2.178}$$

The results reached in (2.177) and (2.178) may be summarized as follows:

$$\boxed{|\psi'\rangle = \hat{U}\,|\psi\rangle, \qquad \langle\psi'| = \langle\psi|\,\hat{U}^\dagger, \qquad \hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger,} \tag{2.179}$$

$$\boxed{|\psi\rangle = \hat{U}^\dagger\,|\psi'\rangle, \qquad \langle\psi| = \langle\psi'|\,\hat{U}, \qquad \hat{A} = \hat{U}^\dagger\hat{A}'\hat{U}.} \tag{2.180}$$

Properties of unitary transformations

- If an operator $\hat{A}$ is Hermitian, its transformed operator $\hat{A}'$ is also Hermitian, since

$$\hat{A}'^\dagger = (\hat{U}\hat{A}\hat{U}^\dagger)^\dagger = \hat{U}\hat{A}^\dagger\hat{U}^\dagger = \hat{U}\hat{A}\hat{U}^\dagger = \hat{A}'. \tag{2.181}$$

- The eigenvalues of $\hat{A}$ and those of its transformed $\hat{A}'$ are the same:

$$\hat{A}\,|\psi_n\rangle = a_n\,|\psi_n\rangle \implies \hat{A}'\,|\psi'_n\rangle = a_n\,|\psi'_n\rangle, \tag{2.182}$$

since

$$\begin{aligned} \hat{A}'\,|\psi'_n\rangle &= (\hat{U}\hat{A}\hat{U}^\dagger)(\hat{U}\,|\psi_n\rangle) = \hat{U}\hat{A}(\hat{U}^\dagger\hat{U})\,|\psi_n\rangle \\ &= \hat{U}\hat{A}\,|\psi_n\rangle = a_n(\hat{U}\,|\psi_n\rangle) = a_n\,|\psi'_n\rangle. \end{aligned} \tag{2.183}$$

- Commutators that are equal to (complex) numbers remain unchanged under unitary transformations, since the transformation of $[\hat{A},\hat{B}] = a$, where a is a complex number, is given by

$$\begin{aligned} [\hat{A}',\hat{B}'] &= [\hat{U}\hat{A}\hat{U}^\dagger,\hat{U}\hat{B}\hat{U}^\dagger] = (\hat{U}\hat{A}\hat{U}^\dagger)(\hat{U}\hat{B}\hat{U}^\dagger) - (\hat{U}\hat{B}\hat{U}^\dagger)(\hat{U}\hat{A}\hat{U}^\dagger) \\ &= \hat{U}[\hat{A},\hat{B}]\hat{U}^\dagger = \hat{U}a\hat{U}^\dagger = a\hat{U}\hat{U}^\dagger = a \\ &= [\hat{A},\hat{B}]. \end{aligned} \tag{2.184}$$

- We can also verify the following general relations:

$$\hat{A} = \beta\hat{B} + \gamma\hat{C} \implies \hat{A}' = \beta\hat{B}' + \gamma\hat{C}', \tag{2.185}$$

$$\hat{A} = \alpha\hat{B}\hat{C}\hat{D} \implies \hat{A}' = \alpha\hat{B}'\hat{C}'\hat{D}', \tag{2.186}$$

where $\hat{A}'$, $\hat{B}'$, $\hat{C}'$, and $\hat{D}'$ are the transforms of $\hat{A}$, $\hat{B}$, $\hat{C}$, and $\hat{D}$, respectively.

- Since the result (2.184) is valid for any complex number, we can state that complex numbers, such as $\langle\psi\,|\,\hat{A}\,|\,\chi\rangle$, remain unchanged under unitary transformations, since

$$\langle\psi'\,|\,\hat{A}'\,|\,\chi'\rangle = (\langle\psi|\,\hat{U}^\dagger)(\hat{U}\hat{A}\hat{U}^\dagger)(\hat{U}\,|\chi\rangle) = \langle\psi\,|\,(\hat{U}^\dagger\hat{U})\hat{A}(\hat{U}^\dagger U)\,|\,\chi\rangle = \langle\psi\,|\,\hat{A}\,|\,\chi\rangle. \tag{2.187}$$

Taking $\hat{A} = \hat{I}$ we see that scalar products of the type

$$\langle\psi'\,|\,\chi'\rangle = \langle\psi\,|\,\chi\rangle \tag{2.188}$$

are invariant under unitary transformations; notably, the norm of a state vector is conserved:

$$\langle\psi'\,|\,\psi'\rangle = \langle\psi\,|\,\psi\rangle. \tag{2.189}$$

- We can also verify that $\left(\hat{U}\hat{A}\hat{U}^\dagger\right)^n = \hat{U}\hat{A}^n\hat{U}^\dagger$ since

$$\begin{aligned}\left(\hat{U}\hat{A}\hat{U}^\dagger\right)^n &= \left(\hat{U}\hat{A}\hat{U}^\dagger\right)\left(\hat{U}\hat{A}\hat{U}^\dagger\right)\cdots\left(\hat{U}\hat{A}\hat{U}^\dagger\right) = \hat{U}\hat{A}(\hat{U}^\dagger\hat{U})\,\hat{A}(\hat{U}^\dagger\hat{U})\cdots(\hat{U}^\dagger\hat{U})\,\hat{A}\hat{U}^\dagger \\ &= \hat{U}\hat{A}^n\hat{U}^\dagger. \end{aligned} \tag{2.190}$$

- We can generalize the previous result to obtain the transformation of any operator function $f(\hat{A})$:

$$\hat{U}f(\hat{A})\hat{U}^\dagger = f(\hat{U}\hat{A}\hat{U}^\dagger) = f(\hat{A}'), \tag{2.191}$$

or more generally

$$\hat{U}f(\hat{A},\hat{B},\hat{C},\cdots)\hat{U}^\dagger = f(\hat{U}\hat{A}\hat{U}^\dagger,\hat{U}\hat{B}\hat{U}^\dagger,\hat{U}\hat{C}\hat{U}^\dagger,\cdots) = f(\hat{A}',\hat{B}',\hat{C}',\cdots). \tag{2.192}$$

A unitary transformation does not change the physics of a system; it merely transforms one description of the system to another physically equivalent description.

In what follows we want to consider two types of unitary transformations: infinitesimal transformations and finite transformations.

2.4.10.2 Infinitesimal Unitary Transformations

Consider an operator $\hat{U}$ which depends on an infinitesimally small real parameter ε and which varies only slightly from the unit operator $\hat{I}$:

$$\hat{U}_\varepsilon(\hat{G}) = \hat{I} + i\varepsilon\hat{G}, \tag{2.193}$$

where $\hat{G}$ is called the *generator* of the infinitesimal transformation. Clearly, $\hat{U}_\varepsilon$ is approximately a unitary transformation only when the parameter ε is real and $\hat{G}$ is Hermitian, since

$$\hat{U}_\varepsilon\hat{U}_\varepsilon^\dagger = (\hat{I} + i\varepsilon\hat{G})(\hat{I} - i\varepsilon\hat{G}^\dagger) \simeq \hat{I} + i\varepsilon(\hat{G} - \hat{G}^\dagger) = \hat{I}, \tag{2.194}$$

where we have neglected the quadratic terms in ε.

The transformation of a state vector $|\psi\rangle$ is

$$|\psi'\rangle = (\hat{I} + i\varepsilon\hat{G})\,|\psi\rangle = |\psi\rangle + \delta|\psi\rangle, \tag{2.195}$$

where

$$\delta|\psi\rangle = i\varepsilon\hat{G}\,|\psi\rangle. \tag{2.196}$$

The transformation of an operator $\hat{A}$ is given by

$$\boxed{\hat{A}' = (\hat{I} + i\varepsilon\hat{G})\hat{A}(\hat{I} - i\varepsilon\hat{G}) \simeq \hat{A} + i\varepsilon[\hat{G},\hat{A}].} \tag{2.197}$$

If $\hat{G}$ commutes with $\hat{A}$, the unitary transformation will leave $\hat{A}$ unchanged, $\hat{A}' = \hat{A}$:

$$[\hat{G},\hat{A}] = 0 \implies \hat{A}' = (\hat{I} + i\varepsilon\hat{G})\hat{A}(\hat{I} - i\varepsilon\hat{G}) = \hat{A}. \tag{2.198}$$

2.4.10.3 Finite Unitary Transformations

We can construct a *finite* unitary transformation from (2.193) by performing a succession of infinitesimal transformations in steps of ε; the application of a series of successive unitary transformations is equivalent to the application of a single unitary transformation. Denoting $\varepsilon = \alpha/N$, where N is an integer and α is a finite parameter, we can apply the same unitary transformation N times; in the limit $N \to +\infty$ we obtain

$$\hat{U}_\alpha(\hat{G}) = \lim_{N\to\infty} \prod_{k=1}^{N} \left(1 + i\frac{\alpha}{N}\hat{G}\right) = \lim_{N\to+\infty} \left(1 + i\frac{\alpha}{N}\hat{G}\right)^N = e^{i\alpha\hat{G}}, \tag{2.199}$$

where $\hat{G}$ is now the generator of the finite transformation and α is its parameter.

As shown in (2.138), $\hat{U}$ is unitary only when the parameter α is real and $\hat{G}$ is Hermitian, since

$$(e^{i\alpha\hat{G}})^\dagger = e^{-i\alpha\hat{G}} = (e^{i\alpha\hat{G}})^{-1}. \tag{2.200}$$

Using the commutation relation (2.130), we can write the transformation $\hat{A}'$ of an operator $\hat{A}$ as follows:

$$\boxed{e^{i\alpha\hat{G}}\hat{A}e^{-i\alpha\hat{G}} = \hat{A} + i\alpha[\hat{G},\hat{A}] + \frac{(i\alpha)^2}{2!}\left[\hat{G},[\hat{G},\hat{A}]\right] + \frac{(i\alpha)^3}{3!}\left[\hat{G},[\hat{G},[\hat{G},\hat{A}]]\right] + \cdots.} \tag{2.201}$$

If $\hat{G}$ commutes with $\hat{A}$, the unitary transformation will leave $\hat{A}$ unchanged, $\hat{A}' = \hat{A}$:

$$[\hat{G},\hat{A}] = 0 \Longrightarrow \hat{A}' = e^{i\alpha\hat{G}}\hat{A}e^{-i\alpha\hat{G}} = \hat{A}. \tag{2.202}$$

In Chapter 3, we will consider some important applications of infinitesimal unitary transformations to study time translations, space translations, space rotations, and conservation laws.

2.5 Representation in Discrete Bases

By analogy with the expansion of Euclidean space vectors in terms of the basis vectors, we need to express any ket $|\psi\rangle$ of the Hilbert space in terms of a complete set of mutually orthonormal base kets. State vectors are then represented by their components in this basis.

2.5.1 Matrix Representation of Kets, Bras, and Operators

Consider a discrete, complete, and orthonormal basis which is made of an infinite[5] set of kets $|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle, \ldots, |\phi_n\rangle$ and denote it by $\{|\phi_n\rangle\}$. Note that the basis $\{|\phi_n\rangle\}$ is discrete, yet it has an infinite number of unit vectors. In the limit $n \to \infty$, the ordering index n of the unit vectors $|\phi_n\rangle$ is *discrete* or *countable*; that is, the sequence $|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle, \ldots$ is countably infinite. As an illustration, consider the special functions, such as the Hermite, Legendre, or Laguerre polynomials, $H_n(x)$, $P_n(x)$, and $L_n(x)$. These polynomials are identified by a discrete index n and by a continuous variable x; although n varies discretely, it can be infinite.

[5]Kets are elements of the Hilbert space, and the dimension of a Hilbert space is infinite.

In Section 2.6, we will consider bases that have a continuous and infinite number of base vectors; in these bases the index n increases continuously. Thus, each basis has a continuum of base vectors.

In this section the notation $\{|\,\phi_n\rangle\}$ will be used to abbreviate an infinitely countable set of vectors (i.e., $|\,\phi_1\rangle, |\,\phi_2\rangle, |\,\phi_3\rangle, \ldots$) of the Hilbert space $\mathcal{H}$. The orthonormality condition of the base kets is expressed by

$$\langle\phi_n \mid \phi_m\rangle = \delta_{nm}, \tag{2.203}$$

where δ_{nm} is the *Kronecker delta* symbol defined by

$$\delta_{nm} = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases} \tag{2.204}$$

The completeness, or closure, relation for this basis is given by

$$\sum_{n=1}^{\infty} |\,\phi_n\rangle\langle\phi_n\,| = \hat{I}, \tag{2.205}$$

where $\hat{I}$ is the unit operator; when the unit operator acts on any ket, it leaves the ket unchanged.

2.5.1.1 Matrix Representation of Kets and Bras

Let us now examine how to represent the vector $|\,\psi\rangle$ within the context of the basis $\{|\,\phi_n\rangle\}$. The completeness property of this basis enables us to expand any state vector $|\,\psi\rangle$ in terms of the base kets $|\,\phi_n\rangle$:

$$|\,\psi\rangle = \hat{I}\,|\,\psi\rangle = \left(\sum_{n=1}^{\infty} |\,\phi_n\rangle\langle\phi_n\,|\right)|\,\psi\rangle = \sum_{n=1}^{\infty} a_n\,|\,\phi_n\rangle, \tag{2.206}$$

where the coefficient a_n, which is equal to $\langle\phi_n \mid \psi\rangle$, represents the projection of $|\,\psi\rangle$ onto $|\,\phi_n\rangle$; a_n is the component of $|\,\psi\rangle$ along the vector $|\,\phi_n\rangle$. Recall that the coefficients a_n are complex numbers. So, within the basis $\{|\,\phi_n\rangle\}$, the ket $|\,\psi\rangle$ is represented by the set of its components, $a_1, a_2, a_3, \ldots$ along $|\,\phi_1\rangle, |\,\phi_2\rangle, |\,\phi_3\rangle, \ldots$, respectively. Hence $|\,\psi\rangle$ can be represented by a *column* vector which has a countably infinite number of components:

$$|\,\psi\rangle \longrightarrow \begin{pmatrix} \langle\phi_1 \mid \psi\rangle \\ \langle\phi_2 \mid \psi\rangle \\ \vdots \\ \langle\phi_n \mid \psi\rangle \\ \vdots \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \vdots \end{pmatrix} \tag{2.207}$$

The bra $\langle\psi\,|$ can be represented by a *row* vector:

$$\begin{aligned} \langle\psi\,| \longrightarrow\ & (\langle\psi \mid \phi_1\rangle\ \langle\psi \mid \phi_2\rangle\ \cdots\ \langle\psi \mid \phi_n\rangle\ \cdots) \\ =\ & (\langle\phi_1 \mid \psi\rangle^*\ \langle\phi_2 \mid \psi\rangle^*\ \cdots\ \langle\phi_n \mid \psi\rangle^*\ \cdots) \\ =\ & (a_1^*\ a_2^*\ \cdots\ a_n^*\ \cdots). \end{aligned} \tag{2.208}$$

Using this representation, we see that a bra-ket $\langle\psi \mid \phi\rangle$ is a complex number equal to the matrix product of the row matrix corresponding to the bra $\langle\psi \mid$ with the column matrix corresponding to the ket $\mid \phi\rangle$:

$$\langle\psi \mid \phi\rangle = (a_1^* \; a_2^* \; \cdots \; a_n^* \; \cdots) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \\ \vdots \end{pmatrix} = \sum_n a_n^* b_n, \tag{2.209}$$

where $b_n = \langle\phi_n \mid \phi\rangle$. We see that, within this representation, the matrices representing $\mid \psi\rangle$ and $\langle\psi \mid$ are Hermitian adjoints of each other.

Remark

A ket $|\psi\rangle$ is normalized if $\langle\psi \mid \psi\rangle = \sum_n |a_n|^2 = 1$. If $|\psi\rangle$ is not normalized and we want to normalize it, we need simply to multiply it by a constant α so that $\langle\alpha\psi \mid \alpha\psi\rangle = |\alpha|^2\langle\psi \mid \psi\rangle = 1$, and hence $\alpha = 1/\sqrt{\langle\psi \mid \psi\rangle}$.

Example 2.14

Consider the following two kets:

$$\mid \psi\rangle = \begin{pmatrix} 5i \\ 2 \\ -i \end{pmatrix}, \qquad \mid \phi\rangle = \begin{pmatrix} 3 \\ 8i \\ -9i \end{pmatrix}.$$

(a) Find $\mid \psi\rangle^*$ and $\langle\psi \mid$.
(b) Is $\mid \psi\rangle$ normalized? If not, normalize it.
(c) Are $\mid \psi\rangle$ and $\mid \phi\rangle$ orthogonal?

Solution

(a) The expressions of $\mid \psi\rangle^*$ and $\langle\psi \mid$ are given by

$$\mid \psi\rangle^* = \begin{pmatrix} -5i \\ 2 \\ i \end{pmatrix}, \qquad \langle\psi \mid \; = \; (-5i \quad 2 \quad i), \tag{2.210}$$

where we have used the fact that $\langle\psi \mid$ is equal to the complex conjugate of the transpose of the ket $\mid \psi\rangle$. Hence, we should reiterate the important fact that $\mid \psi\rangle^* \neq \langle\psi \mid$.

(b) The norm of $\mid \psi\rangle$ is given by

$$\langle\psi \mid \psi\rangle = (-5i \quad 2 \quad i) \begin{pmatrix} 5i \\ 2 \\ -i \end{pmatrix} = (-5i)(5i) + (2)(2) + (i)(-i) = 30. \tag{2.211}$$

Thus, $\mid \psi\rangle$ is not normalized. By multiplying it with $1/\sqrt{30}$, it becomes normalized:

$$\mid \chi\rangle = \frac{1}{\sqrt{30}} \mid \psi\rangle = \frac{1}{\sqrt{30}} \begin{pmatrix} 5i \\ 2 \\ -i \end{pmatrix} \quad \Longrightarrow \quad \langle\chi \mid \chi\rangle = 1. \tag{2.212}$$

(c) The kets $|\psi\rangle$ and $|\phi\rangle$ are not orthogonal since their scalar product is not zero:

$$\langle\psi|\phi\rangle = \begin{pmatrix} -5i & 2 & i \end{pmatrix}\begin{pmatrix} 3 \\ 8i \\ -9i \end{pmatrix} = (-5i)(3) + (2)(8i) + (i)(-9i) = 9 + i. \tag{2.213}$$

2.5.1.2 Matrix Representation of Operators

For each linear operator $\hat{A}$, we can write

$$\hat{A} = \hat{I}\hat{A}\hat{I} = \left(\sum_{n=1}^{\infty} |\phi_n\rangle\langle\phi_n|\right)\hat{A}\left(\sum_{m=1}^{\infty} |\phi_m\rangle\langle\phi_m|\right) = \sum_{nm} A_{nm}\,|\phi_n\rangle\langle\phi_m|, \tag{2.214}$$

where A_{nm} is the nm matrix element of the operator $\hat{A}$:

$$A_{nm} = \langle\phi_n|\hat{A}|\phi_m\rangle. \tag{2.215}$$

We see that the operator $\hat{A}$ is represented, within the basis $\{|\phi_n\rangle\}$, by a *square* matrix A (A without a hat designates a matrix), which has a countably infinite number of columns and a countably infinite number of rows:

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2.216}$$

For instance, the *unit* operator $\hat{I}$ is represented by the unit matrix; when the unit matrix is multiplied with another matrix, it leaves that unchanged:

$$I = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.217}$$

In summary, kets are represented by column vectors, bras by row vectors, bra-kets by numbers (either real or complex numbers), and operators by square matrices.

2.5.1.3 Matrix Representation of Some Other Operators

(a) Hermitian adjoint operation

Let us now look at the matrix representation of the Hermitian adjoint operation of an operator. First, recall that the *transpose* of a matrix A, denoted by A^T, is obtained by interchanging the rows with the columns:

$$(A^T)_{nm} = A_{mn} \quad \text{or} \quad \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}^T = \begin{pmatrix} A_{11} & A_{21} & A_{31} & \cdots \\ A_{12} & A_{22} & A_{32} & \cdots \\ A_{13} & A_{23} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.218}$$

Similarly, the transpose of a column matrix is a row matrix, and the transpose of a row matrix is a column matrix:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \vdots \end{pmatrix}^T = \begin{pmatrix} a_1 & a_2 & \cdots & a_n & \cdots \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a_1 & a_2 & \cdots & a_n & \cdots \end{pmatrix}^T = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \vdots \end{pmatrix}. \tag{2.219}$$

So a square matrix A is symmetric if it is equal to its transpose, $A^T = A$. A skew-symmetric matrix is a square matrix whose transpose equals the negative of the matrix, $A^T = -A$.

The *complex conjugate* of a matrix is obtained by simply taking the complex conjugate of all its elements: $(A^*)_{nm} = (A_{nm})^*$.

The matrix which represents the operator $\hat{A}^\dagger$ is obtained by taking the complex conjugate of the matrix transpose of A:

$$A^\dagger = (A^T)^* \quad \text{or} \quad (\hat{A}^\dagger)_{nm} = \langle \phi_n \mid \hat{A}^\dagger \mid \phi_m \rangle = \langle \phi_m \mid \hat{A} \mid \phi_n \rangle^* = A^*_{mn}; \tag{2.220}$$

that is,

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}^\dagger = \begin{pmatrix} A^*_{11} & A^*_{21} & A^*_{31} & \cdots \\ A^*_{12} & A^*_{22} & A^*_{32} & \cdots \\ A^*_{13} & A^*_{23} & A^*_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.221}$$

If an operator $\hat{A}$ is Hermitian, its matrix satisfies this condition:

$$(A^T)^* = A \quad \text{or} \quad A^*_{mn} = A_{nm}. \tag{2.222}$$

The diagonal elements of a Hermitian matrix therefore must be real numbers. Note that *a Hermitian matrix must be square.*

(b) Inverse and unitary operators

A matrix has an inverse only if it is square and its determinant is nonzero; a matrix that has an inverse is called a nonsingular matrix and a matrix that has no inverse is called a singular

matrix. The elements A^{-1}_{nm} of the inverse matrix A^{-1}, representing an operator $\hat{A}^{-1}$, are given by the relation

$$A^{-1}_{nm} = \frac{\text{cofactor of } A_{mn}}{\text{determinant of } A} \quad \text{or} \quad A^{-1} = \frac{B^T}{\text{determinant of } A}, \tag{2.223}$$

where B is the matrix of *cofactors* (also called the minor); the cofactor of element A_{mn} is equal to $(-1)^{m+n}$ times the determinant of the submatrix obtained from A by removing the mth row and the nth column. Note that when the matrix, representing an operator, has a determinant equal to zero, this operator does not possess an inverse. Note that $A^{-1}A = AA^{-1} = I$ where I is the unit matrix.

The inverse of a product of matrices is obtained as follows:

$$(ABC\cdots PQ)^{-1} = Q^{-1}P^{-1}\cdots C^{-1}B^{-1}A^{-1}. \tag{2.224}$$

The inverse of the inverse of a matrix is equal to the matrix itself, $\left(A^{-1}\right)^{-1} = A$.

A *unitary* operator $\hat{U}$ is represented by a unitary matrix. A matrix U is said to be unitary if its inverse is equal to its adjoint:

$$U^{-1} = U^\dagger \quad \text{or} \quad U^\dagger U = I, \tag{2.225}$$

where I is the unit matrix.

Example 2.15 (Inverse of a matrix)

Calculate the inverse of the matrix $A = \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix}$. Is this matrix unitary?

Solution

Since the determinant of A is $\det(A) = -4 + 16i$, we have $A^{-1} = B^T/(-4 + 16i)$, where the elements of the cofactor matrix B are given by $B_{nm} = (-1)^{n+m}$ times the determinant of the submatrix obtained from A by removing the nth row and the mth column. In this way, we have

$$B_{11} = (-1)^{1+1}\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} = (-1)^2\begin{vmatrix} 1 & 5 \\ -i & -2 \end{vmatrix} = -2 + 5i, \tag{2.226}$$

$$B_{12} = (-1)^{1+2}\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} = (-1)^3\begin{vmatrix} 3 & 5 \\ 0 & -2 \end{vmatrix} = 6, \tag{2.227}$$

$$B_{13} = (-1)^{1+3}\begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} = (-1)^4\begin{vmatrix} 3 & 1 \\ 0 & -i \end{vmatrix} = -3i, \tag{2.228}$$

$$B_{21} = (-1)^3\begin{vmatrix} i & 0 \\ -i & -2 \end{vmatrix} = 2i, \qquad B_{22} = (-1)^4\begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix} = -4, \tag{2.229}$$

$$B_{23} = (-1)^5\begin{vmatrix} 2 & i \\ 0 & -i \end{vmatrix} = 2i, \qquad B_{31} = (-1)^4\begin{vmatrix} i & 0 \\ 1 & 5 \end{vmatrix} = 5i, \tag{2.230}$$

$$B_{32} = (-1)^5\begin{vmatrix} 2 & 0 \\ 3 & 5 \end{vmatrix} = -10, \qquad B_{33} = (-1)^6\begin{vmatrix} 2 & i \\ 3 & 1 \end{vmatrix} = 2 - 3i, \tag{2.231}$$

and hence

$$B = \begin{pmatrix} -2+5i & 6 & -3i \\ 2i & -4 & 2i \\ 5i & -10 & 2-3i \end{pmatrix}. \tag{2.232}$$

Taking the transpose of B, we obtain

$$\begin{aligned} A^{-1} &= \frac{1}{-4+16i} B^T = \frac{-1-4i}{68} \begin{pmatrix} -2+5i & 2i & 5i \\ 6 & -4 & -10 \\ -3i & 2i & 2-3i \end{pmatrix} \\ &= \frac{1}{68} \begin{pmatrix} 22+3i & 8-2i & 20-5i \\ -6-24i & 4+16i & 10+40i \\ -12+3i & 8-2i & -14-5i \end{pmatrix}. \end{aligned} \tag{2.233}$$

Clearly, this matrix is not unitary since its inverse is not equal to its Hermitian adjoint: $A^{-1} \neq A^\dagger$.

(c) Matrix representation of $|\psi\rangle\langle\psi|$

It is now easy to see that the product $|\psi\rangle\langle\psi|$ is indeed an operator, since its representation within $\{|\phi_n\rangle\}$ is a square matrix:

$$|\psi\rangle\langle\psi| = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} (a_1^* \;\; a_2^* \;\; a_3^* \;\; \cdots) = \begin{pmatrix} a_1a_1^* & a_1a_2^* & a_1a_3^* & \cdots \\ a_2a_1^* & a_2a_2^* & a_2a_3^* & \cdots \\ a_3a_1^* & a_3a_2^* & a_3a_3^* & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.234}$$

(d) Trace of an operator

The trace $\mathrm{Tr}(\hat{A})$ of an operator $\hat{A}$ is given, within an orthonormal basis $\{|\phi_n\rangle\}$, by the expression

$$\mathrm{Tr}(\hat{A}) = \sum_n \langle\phi_n|\hat{A}|\phi_n\rangle = \sum_n A_{nn}; \tag{2.235}$$

we will see later that the trace of an operator does not depend on the basis. The trace of a matrix is equal to the sum of its diagonal elements:

$$\mathrm{Tr}\begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = A_{11} + A_{22} + A_{33} + \cdots. \tag{2.236}$$

Properties of the trace

We can ascertain that

$$\mathrm{Tr}(\hat{A}^\dagger) = (\mathrm{Tr}(\hat{A}))^*, \tag{2.237}$$

$$\mathrm{Tr}(\alpha\hat{A} + \beta\hat{B} + \gamma\hat{C} + \cdots) = \alpha\mathrm{Tr}(\hat{A}) + \beta\mathrm{Tr}(\hat{B}) + \gamma\mathrm{Tr}(\hat{C}) + \cdots, \tag{2.238}$$

and that the trace of a product of operators is invariant under their cyclic permutations:

$$\mathrm{Tr}(\hat{A}\hat{B}\hat{C}\hat{D}\hat{E}) = \mathrm{Tr}(\hat{E}\hat{A}\hat{B}\hat{C}\hat{D}) = \mathrm{Tr}(\hat{D}\hat{E}\hat{A}\hat{B}\hat{C}) = \mathrm{Tr}(\hat{C}\hat{D}\hat{E}\hat{A}\hat{B}) = \cdots. \tag{2.239}$$

Additionally, for any two operators $\hat{A}$ and $\hat{B}$, all commutators have zero traces

$$\mathrm{Tr}\left([\hat{A}, \hat{B}]\right) = 0. \tag{2.240}$$

Example 2.16

(a) Show that $\mathrm{Tr}(\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{A})$.

(b) Show that the trace of a commutator is always zero.

(c) Illustrate the results shown in (a) and (b) on the following matrices:

$$A = \begin{pmatrix} 8-2i & 4i & 0 \\ 1 & 0 & 1-i \\ -8 & i & 6i \end{pmatrix}, \qquad B = \begin{pmatrix} -i & 2 & 1-i \\ 6 & 1+i & 3i \\ 1 & 5+7i & 0 \end{pmatrix}.$$

Solution

(a) Using the definition of the trace,

$$\mathrm{Tr}(\hat{A}\hat{B}) = \sum_n \langle \phi_n \mid \hat{A}\hat{B} \mid \phi_n \rangle, \tag{2.241}$$

and inserting the unit operator between $\hat{A}$ and $\hat{B}$ we have

$$\begin{aligned} \mathrm{Tr}(\hat{A}\hat{B}) &= \sum_n \langle \phi_n \mid \hat{A} \left(\sum_m \mid \phi_m \rangle\langle \phi_m \mid \right) \hat{B} \mid \phi_n \rangle = \sum_{nm} \langle \phi_n \mid \hat{A} \mid \phi_m \rangle \langle \phi_m \mid \hat{B} \mid \phi_n \rangle \\ &= \sum_{nm} A_{nm} B_{mn}. \end{aligned} \tag{2.242}$$

On the other hand, since $\mathrm{Tr}(\hat{A}\hat{B}) = \sum_n \langle \phi_n \mid \hat{A}\hat{B} \mid \phi_n \rangle$, we have

$$\begin{aligned} \mathrm{Tr}(\hat{B}\hat{A}) &= \sum_m \langle \phi_m \mid \hat{B} \left(\sum_n \mid \phi_n \rangle\langle \phi_n \mid \right) \hat{A} \mid \phi_m \rangle = \sum_m \langle \phi_m \mid \hat{B} \mid \phi_n \rangle \langle \phi_n \mid \hat{A} \mid \phi_m \rangle \\ &= \sum_{nm} B_{mn} A_{nm}. \end{aligned} \tag{2.243}$$

Comparing (2.242) and (2.243), we see that $\mathrm{Tr}(\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{A})$.

(b) Since $\mathrm{Tr}(\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{A})$ we can infer at once that the trace of any commutator is always zero:

$$\mathrm{Tr}([\hat{A}, \hat{B}]) = \mathrm{Tr}(\hat{A}\hat{B}) - \mathrm{Tr}(\hat{B}\hat{A}) = 0. \tag{2.244}$$

(c) Let us verify that the traces of the products AB and BA are equal. Since

$$AB = \begin{pmatrix} -2+16i & 12 & -6-10i \\ 1-2i & 14+2i & 1-i \\ 20i & -59+31i & -11+8i \end{pmatrix}, \quad BA = \begin{pmatrix} -8 & 5+i & 8+4i \\ 49-35i & -3+24i & -16 \\ 13+5i & 4i & 12+2i \end{pmatrix}, \tag{2.245}$$

we have

$$\mathrm{Tr}(AB) = \mathrm{Tr} \begin{pmatrix} -2+16i & 12 & -6-10i \\ 1-2i & 14+2i & 1-i \\ 20i & -59+31i & -11+8i \end{pmatrix} = 1 + 26i, \tag{2.246}$$

$$\mathrm{Tr}(BA) = \mathrm{Tr}\begin{pmatrix} -8 & 5+i & 8+4i \\ 49-35i & -3+24i & -16 \\ 13+5i & 4i & 12+2i \end{pmatrix} = 1+26i = \mathrm{Tr}(AB). \tag{2.247}$$

This leads to $\mathrm{Tr}(AB) - \mathrm{Tr}(BA) = (1+26i) - (1+26i) = 0$ or $\mathrm{Tr}([A, B]) = 0$.

2.5.1.4 Matrix Representation of Several Other Quantities

(a) Matrix representation of $|\phi\rangle = \hat{A}\,|\psi\rangle$

The relation $|\phi\rangle = \hat{A}\,|\psi\rangle$ can be cast into the algebraic form $\hat{I}\,|\phi\rangle = \hat{I}\hat{A}\hat{I}\,|\psi\rangle$ or

$$\left(\sum_n |\phi_n\rangle\langle\phi_n|\right)|\phi\rangle = \left(\sum_n |\phi_n\rangle\langle\phi_n|\right)\hat{A}\left(\sum_m |\phi_m\rangle\langle\phi_m|\right)|\psi\rangle, \tag{2.248}$$

which in turn can be written as

$$\sum_n b_n\,|\phi_n\rangle = \sum_{nm} a_m\,|\phi_n\rangle\langle\phi_n|\hat{A}|\phi_m\rangle = \sum_{nm} a_m A_{nm}\,|\phi_n\rangle, \tag{2.249}$$

where $b_n = \langle\phi_n|\phi\rangle$, $A_{nm} = \langle\phi_n|\hat{A}|\phi_m\rangle$, and $a_m = \langle\phi_m|\psi\rangle$. It is easy to see that (2.249) yields $b_n = \sum_m A_{nm}a_m$; hence the matrix representation of $|\phi\rangle = \hat{A}\,|\psi\rangle$ is given by

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix}. \tag{2.250}$$

(b) Matrix representation of $\langle\phi|\hat{A}|\psi\rangle$

As for $\langle\phi|\hat{A}|\psi\rangle$ we have

$$\begin{aligned} \langle\phi|\hat{A}|\psi\rangle &= \langle\phi|\hat{I}\hat{A}\hat{I}|\psi\rangle = \langle\phi|\left(\sum_{n=1}^{\infty}|\phi_n\rangle\langle\phi_n|\right)\hat{A}\left(\sum_{m=1}^{\infty}|\phi_m\rangle\langle\phi_m|\right)|\psi\rangle \\ &= \sum_{nm}\langle\phi|\phi_n\rangle\langle\phi_n|\hat{A}|\phi_m\rangle\langle\phi_m|\psi\rangle \\ &= \sum_{nm} b_n^* A_{nm} a_m. \end{aligned} \tag{2.251}$$

This is a complex number; its matrix representation goes as follows:

$$\langle\phi|\hat{A}|\psi\rangle \longrightarrow (b_1^* \;\; b_2^* \;\; b_3^* \;\; \cdots)\begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix}. \tag{2.252}$$

Remark
It is now easy to see explicitly why products of the type $|\psi\rangle\,|\phi\rangle$, $\langle\psi|\,\langle\phi|$, $\hat{A}\langle\psi|$, or $|\psi\rangle\hat{A}$ are forbidden. They cannot have matrix representations; they are nonsensical. For instance, $|\psi\rangle\,|\phi\rangle$ is represented by the product of two column matrices:

$$|\psi\rangle\,|\phi\rangle \longrightarrow \begin{pmatrix} \langle\phi_1|\psi\rangle \\ \langle\phi_2|\psi\rangle \\ \vdots \end{pmatrix} \begin{pmatrix} \langle\phi_1|\phi\rangle \\ \langle\phi_2|\phi\rangle \\ \vdots \end{pmatrix}. \tag{2.253}$$

This product is clearly not possible to perform, for the product of two matrices is possible only when the number of columns of the first is equal to the number of rows of the second; in (2.253) the first matrix has one single column and the second an infinite number of rows.

2.5.1.5 Properties of a Matrix A

- Real if $A = A^*$ or $A_{mn} = A^*_{mn}$
- Imaginary if $A = -A^*$ or $A_{mn} = -A^*_{mn}$
- Symmetric if $A = A^T$ or $A_{mn} = A_{nm}$
- Antisymmetric if $A = -A^T$ or $A_{mn} = -A_{nm}$ with $A_{mm} = 0$
- Hermitian if $A = A^\dagger$ or $A_{mn} = A^*_{nm}$
- Anti-Hermitian if $A = -A^\dagger$ or $A_{mn} = -A^*_{nm}$
- Orthogonal if $A^T = A^{-1}$ or $AA^T = I$ or $(AA^T)_{mn} = \delta_{mn}$
- Unitary if $A^\dagger = A^{-1}$ or $AA^\dagger = I$ or $(AA^\dagger)_{mn} = \delta_{mn}$

Example 2.17
Consider a matrix A (which represents an operator $\hat{A}$), a ket $|\psi\rangle$, and a bra $\langle\phi|$:

$$A = \begin{pmatrix} 5 & 3+2i & 3i \\ -i & 3i & 8 \\ 1-i & 1 & 4 \end{pmatrix}, \quad |\psi\rangle = \begin{pmatrix} -1+i \\ 3 \\ 2+3i \end{pmatrix}, \quad \langle\phi| = \begin{pmatrix} 6 & -i & 5 \end{pmatrix}.$$

(a) Calculate the quantities $A\,|\psi\rangle$, $\langle\phi|\,A$, $\langle\phi|\,A\,|\psi\rangle$, and $|\psi\rangle\langle\phi|$.

(b) Find the complex conjugate, the transpose, and the Hermitian conjugate of A, $|\psi\rangle$, and $\langle\phi|$.

(c) Calculate $\langle\phi|\psi\rangle$ and $\langle\psi|\phi\rangle$; are they equal? Comment on the differences between the complex conjugate, Hermitian conjugate, and transpose of kets and bras.

Solution

(a) The calculations are straightforward:

$$A\,|\,\psi\rangle = \begin{pmatrix} 5 & 3+2i & 3i \\ -i & 3i & 8 \\ 1-i & 1 & 4 \end{pmatrix}\begin{pmatrix} -1+i \\ 3 \\ 2+3i \end{pmatrix} = \begin{pmatrix} -5+17i \\ 17+34i \\ 11+14i \end{pmatrix}, \tag{2.254}$$

$$\langle\phi\,|\,A = \begin{pmatrix} 6 & -i & 5 \end{pmatrix}\begin{pmatrix} 5 & 3+2i & 3i \\ -i & 3i & 8 \\ 1-i & 1 & 4 \end{pmatrix} = \begin{pmatrix} 34-5i & 26+12i & 20+10i \end{pmatrix}, \tag{2.255}$$

$$\langle\phi\,|\,A\,|\,\psi\rangle = \begin{pmatrix} 6 & -i & 5 \end{pmatrix}\begin{pmatrix} 5 & 3+2i & 3i \\ -i & 3i & 8 \\ 1-i & 1 & 4 \end{pmatrix}\begin{pmatrix} -1+i \\ 3 \\ 2+3i \end{pmatrix} = 59+155i, \tag{2.256}$$

$$|\,\psi\rangle\langle\phi\,| = \begin{pmatrix} -1+i \\ 3 \\ 2+3i \end{pmatrix}\begin{pmatrix} 6 & -i & 5 \end{pmatrix} = \begin{pmatrix} -6+6i & 1+i & -5+5i \\ 18 & -3i & 15 \\ 12+18i & 3-2i & 10+15i \end{pmatrix}. \tag{2.257}$$

(b) To obtain the complex conjugate of A, $|\,\psi\rangle$, and $\langle\phi\,|$, we need simply to take the complex conjugate of their elements:

$$A^* = \begin{pmatrix} 5 & 3-2i & -3i \\ i & -3i & 8 \\ 1+i & 1 & 4 \end{pmatrix}, \quad |\,\psi\rangle^* = \begin{pmatrix} -1-i \\ 3 \\ 2-3i \end{pmatrix}, \quad \langle\phi\,|^* = \begin{pmatrix} 6 & i & 5 \end{pmatrix}. \tag{2.258}$$

For the transpose of A, $|\,\psi\rangle$, and $\langle\phi\,|$, we simply interchange columns with rows:

$$A^T = \begin{pmatrix} 5 & -i & 1-i \\ 3+2i & 3i & 1 \\ 3i & 8 & 4 \end{pmatrix}, \quad |\,\psi\rangle^T = \begin{pmatrix} -1+i & 3 & 2+3i \end{pmatrix}, \quad \langle\phi\,|^T = \begin{pmatrix} 6 \\ -i \\ 5 \end{pmatrix}. \tag{2.259}$$

The Hermitian conjugate can be obtained by taking the complex conjugates of the transpose expressions calculated above: $A^\dagger = (A^T)^*$, $|\,\psi\rangle^\dagger = \left(|\,\psi\rangle^T\right)^* = \langle\psi\,|$, $\langle\phi\,|^\dagger = \left(\langle\phi\,|^T\right)^* = |\,\phi\rangle$:

$$A^\dagger = \begin{pmatrix} 5 & i & 1+i \\ 3-2i & -3i & 1 \\ -3i & 8 & 4 \end{pmatrix}, \quad \langle\psi\,| = \begin{pmatrix} -1-i & 3 & 2-3i \end{pmatrix}, \quad |\,\phi\rangle = \begin{pmatrix} 6 \\ i \\ 5 \end{pmatrix}. \tag{2.260}$$

(c) Using the kets and bras above, we can easily calculate the needed scalar products:

$$\langle\phi\,|\,\psi\rangle = \begin{pmatrix} 6 & -i & 5 \end{pmatrix}\begin{pmatrix} -1+i \\ 3 \\ 2+3i \end{pmatrix} = 6(-1+i) + (-i)(3) + 5(2+3i) = 4+18i, \tag{2.261}$$

$$\langle\psi\,|\,\phi\rangle = \begin{pmatrix} -1-i & 3 & 2-3i \end{pmatrix}\begin{pmatrix} 6 \\ i \\ 5 \end{pmatrix} = 6(-1-i) + (i)(3) + 5(2-3i) = 4-18i. \tag{2.262}$$

We see that $\langle\phi\,|\,\psi\rangle$ and $\langle\psi\,|\,\phi\rangle$ are not equal; they are complex conjugates of each other:

$$\langle\psi\,|\,\phi\rangle = \langle\phi\,|\,\psi\rangle^* = 4-18i. \tag{2.263}$$

Remark
We should underscore the importance of the differences between $|\psi\rangle^*$, $|\psi\rangle^T$, and $|\psi\rangle^\dagger$. Most notably, we should note (from equations (2.258)–(2.260)) that $|\psi\rangle^*$ is a ket, while $|\psi\rangle^T$ and $|\psi\rangle^\dagger$ are bras. Additionally, we should note that $\langle\phi|^*$ is a bra, while $\langle\phi|^T$ and $\langle\phi|^\dagger$ are kets.

2.5.2 Change of Bases and Unitary Transformations

In a Euclidean space, a vector $\vec{A}$ may be represented by its components in different coordinate systems or in different bases. The transformation from one basis to the other is called a change of basis. The components of $\vec{A}$ in a given basis can be expressed in terms of the components of $\vec{A}$ in another basis by means of a *transformation* matrix.

Similarly, state vectors and operators of quantum mechanics may also be represented in different bases. In this section we are going to study how to transform from one basis to another. That is, knowing the components of kets, bras, and operators in a basis $\{|\phi_n\rangle\}$, how does one determine the corresponding components in a different basis $\{|\phi'_n\rangle\}$? Assuming that $\{|\phi_n\rangle\}$ and $\{|\phi'_n\rangle\}$ are two different bases, we can expand each ket $|\phi_n\rangle$ of the old basis in terms of the new basis $\{|\phi'_n\rangle\}$ as follows:

$$|\phi_n\rangle = \left(\sum_m |\phi'_m\rangle\langle\phi'_m|\right)|\phi_n\rangle = \sum_m U_{mn}|\phi'_m\rangle, \tag{2.264}$$

where

$$U_{mn} = \langle\phi'_m|\phi_n\rangle. \tag{2.265}$$

The matrix U, providing the transformation from the old basis $\{|\phi_n\rangle\}$ to the new basis $\{|\phi'_n\rangle\}$, is given by

$$U = \begin{pmatrix} \langle\phi'_1|\phi_1\rangle & \langle\phi'_1|\phi_2\rangle & \langle\phi'_1|\phi_3\rangle \\ \langle\phi'_2|\phi_1\rangle & \langle\phi'_2|\phi_2\rangle & \langle\phi'_2|\phi_3\rangle \\ \langle\phi'_3|\phi_1\rangle & \langle\phi'_3|\phi_2\rangle & \langle\phi'_3|\phi_3\rangle \end{pmatrix}. \tag{2.266}$$

Example 2.18 (Unitarity of the transformation matrix)
Let U be a transformation matrix which connects two *complete* and *orthonormal* bases $\{|\phi_n\rangle\}$ and $\{|\phi'_n\rangle\}$. Show that U is *unitary*.

Solution
For this we need to prove that $\hat{U}\hat{U}^\dagger = \hat{I}$, which reduces to showing that $\langle\phi_m|\hat{U}\hat{U}^\dagger|\phi_n\rangle = \delta_{mn}$. This goes as follows:

$$\langle\phi_m|\hat{U}\hat{U}^\dagger|\phi_n\rangle = \langle\phi_m|\hat{U}\left(\sum_l |\phi_l\rangle\langle\phi_l|\right)\hat{U}^\dagger|\phi_n\rangle = \sum_l U_{ml}U^*_{nl}, \tag{2.267}$$

where $U_{ml} = \langle\phi_m|\hat{U}|\phi_l\rangle$ and $U^*_{nl} = \langle\phi_l|\hat{U}^\dagger|\phi_n\rangle = \langle\phi_n|\hat{U}|\phi_l\rangle^*$. According to (2.265), $U_{ml} = \langle\phi'_m|\phi_l\rangle$ and $U^*_{nl} = \langle\phi_l|\phi'_n\rangle$; we can thus rewrite (2.267) as

$$\sum_l U_{ml}U^*_{nl} = \sum_l \langle\phi'_m|\phi_l\rangle\langle\phi_l|\phi'_n\rangle = \langle\phi'_m|\phi'_n\rangle = \delta_{mn}. \tag{2.268}$$

Combining (2.267) and (2.268), we infer $\langle \phi_m \mid \hat{U}\hat{U}^\dagger \mid \phi_n \rangle = \delta_{mn}$, or $\hat{U}\hat{U}^\dagger = \hat{I}$.

2.5.2.1 Transformations of Kets, Bras, and Operators

The components $\langle \phi'_n \mid \psi \rangle$ of a state vector $\mid \psi \rangle$ in a new basis $\{\mid \phi'_n \rangle\}$ can be expressed in terms of the components $\langle \phi_n \mid \psi \rangle$ of $\mid \psi \rangle$ in an old basis $\{\mid \phi_n \rangle\}$ as follows:

$$\langle \phi'_m \mid \psi \rangle = \langle \phi'_m \mid \hat{I} \mid \psi \rangle = \langle \phi'_m \mid \left(\sum_n \mid \phi_n \rangle\langle \phi_n \mid \right) \mid \psi \rangle = \sum_n U_{mn} \langle \phi_n \mid \psi \rangle. \tag{2.269}$$

This relation, along with its complex conjugate, can be generalized into

$$\mid \psi_{new} \rangle = \hat{U} \mid \psi_{old} \rangle, \qquad \langle \psi_{new} \mid = \langle \psi_{old} \mid \hat{U}^\dagger. \tag{2.270}$$

Let us now examine how operators transform when we change from one basis to another. The matrix elements $A'_{mn} = \langle \phi'_m \mid \hat{A} \mid \phi'_n \rangle$ of an operator $\hat{A}$ in the new basis can be expressed in terms of the old matrix elements, $A_{jl} = \langle \phi_j \mid \hat{A} \mid \phi_l \rangle$, as follows:

$$A'_{mn} = \langle \phi'_m \mid \left(\sum_j \mid \phi_j \rangle\langle \phi_j \mid \right) \hat{A} \left(\sum_l \mid \phi_l \rangle\langle \phi_l \mid \right) \mid \phi'_n \rangle = \sum_{jl} U_{mj} A_{jl} U^*_{nl}; \tag{2.271}$$

that is,

$$\hat{A}_{new} = \hat{U}\hat{A}_{old}\hat{U}^\dagger \qquad \text{or} \qquad \hat{A}_{old} = \hat{U}^\dagger \hat{A}_{new} \hat{U}. \tag{2.272}$$

We may summarize the results of the change of basis in the following relations:

$$\boxed{\mid \psi_{new} \rangle = \hat{U} \mid \psi_{old} \rangle, \qquad \langle \psi_{new} \mid = \langle \psi_{old} \mid \hat{U}^\dagger, \qquad \hat{A}_{new} = \hat{U}\hat{A}_{old}\hat{U}^\dagger,} \tag{2.273}$$

or

$$\boxed{\mid \psi_{old} \rangle = \hat{U}^\dagger \mid \psi_{new} \rangle, \qquad \langle \psi_{old} \mid = \langle \psi_{new} \mid \hat{U}, \qquad \hat{A}_{old} = \hat{U}^\dagger \hat{A}_{new} \hat{U}.} \tag{2.274}$$

These relations are similar to the ones we derived when we studied unitary transformations; see (2.179) and (2.180).

Example 2.19
Show that the operator $\hat{U} = \sum_n \mid \phi'_n \rangle\langle \phi_n \mid$ satisfies all the properties discussed above.

Solution
First, note that $\hat{U}$ is unitary:

$$\hat{U}\hat{U}^\dagger = \sum_{nl} \mid \phi'_n \rangle\langle \phi_n \mid \phi_l \rangle\langle \phi'_l \mid = \sum_{nl} \mid \phi'_n \rangle\langle \phi'_l \mid \delta_{nl} = \sum_n \mid \phi'_n \rangle\langle \phi'_n \mid = \hat{I}. \tag{2.275}$$

Second, the action of $\hat{U}$ on a ket of the old basis gives the corresponding ket from the new basis:

$$\hat{U} \mid \phi_m\rangle = \sum_n \mid \phi'_n\rangle\langle\phi_n \mid \phi_m\rangle = \sum_n \mid \phi'_n\rangle\delta_{nm} = \mid \phi'_m\rangle. \tag{2.276}$$

We can also verify that the action $\hat{U}^\dagger$ on a ket of the new basis gives the corresponding ket from the old basis:

$$\hat{U}^\dagger \mid \phi'_m\rangle = \sum_l \mid \phi_l\rangle\langle\phi'_l \mid \phi'_m\rangle = \sum_l \mid \phi_l\rangle\delta_{lm} = \mid \phi_m\rangle. \tag{2.277}$$

How does a trace transform under unitary transformations? Using the cyclic property of the trace, $\mathrm{Tr}(\hat{A}\hat{B}\hat{C}) = \mathrm{Tr}(\hat{C}\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{C}\hat{A})$, we can ascertain that

$$\mathrm{Tr}(\hat{A}') = \mathrm{Tr}(\hat{U}\hat{A}\hat{U}^\dagger) = \mathrm{Tr}(\hat{U}^\dagger\hat{U}\hat{A}) = \mathrm{Tr}(\hat{A}), \tag{2.278}$$

$$\begin{aligned} \mathrm{Tr}\left(\mid \phi_n\rangle\langle\phi_m \mid\right) &= \sum_l \langle\phi_l \mid \phi_n\rangle\langle\phi_m \mid \phi_l\rangle = \sum_l \langle\phi_m \mid \phi_l\rangle\langle\phi_l \mid \phi_n\rangle \\ &= \langle\phi_m \mid \left(\sum_l \mid \phi_l\rangle\langle\phi_l \mid\right) \mid \phi_n\rangle = \langle\phi_m \mid \phi_n\rangle = \delta_{mn}, \end{aligned} \tag{2.279}$$

$$\mathrm{Tr}\left(\mid \phi'_m\rangle\langle\phi_n \mid\right) = \langle\phi_n \mid \phi'_m\rangle. \tag{2.280}$$

Example 2.20 (The trace is base independent)
Show that the trace of an operator does not depend on the basis in which it is expressed.

Solution
Let us show that the trace of an operator $\hat{A}$ in a basis $\{\mid \phi_n\rangle\}$ is equal to its trace in another basis $\{\mid \phi'_n\rangle\}$. First, the trace of $\hat{A}$ in the basis $\{\mid \phi_n\rangle\}$ is given by

$$\mathrm{Tr}(\hat{A}) = \sum_n \langle\phi_n \mid \hat{A} \mid \phi_n\rangle \tag{2.281}$$

and in $\{\mid \phi'_n\rangle\}$ by

$$\mathrm{Tr}(\hat{A}) = \sum_n \langle\phi'_n \mid \hat{A} \mid \phi'_n\rangle. \tag{2.282}$$

Starting from (2.281) and using the completeness of the other basis, $\{\mid \phi'_n\rangle\}$, we have

$$\begin{aligned} \mathrm{Tr}(\hat{A}) &= \sum_n \langle\phi_n \mid \hat{A} \mid \phi_n\rangle = \sum_n \langle\phi_n \mid \left(\sum_m \mid \phi'_m\rangle\langle\phi'_m \mid\right) \hat{A} \mid \phi_n\rangle \\ &= \sum_{nm} \langle\phi_n \mid \phi'_m\rangle\langle\phi'_m \mid \hat{A} \mid \phi_n\rangle. \end{aligned} \tag{2.283}$$

All we need to do now is simply to interchange the positions of the numbers (scalars) $\langle \phi_n \mid \phi'_m \rangle$ and $\langle \phi'_m \mid \hat{A} \mid \phi_n \rangle$:

$$\mathrm{Tr}(\hat{A}) = \sum_m \langle \phi'_m \mid \hat{A} \left(\sum_n \mid \phi_n \rangle \langle \phi_n \mid \right) \mid \phi'_m \rangle = \sum_m \langle \phi'_m \mid \hat{A} \mid \phi'_m \rangle. \tag{2.284}$$

From (2.283) and (2.284) we see that

$$\mathrm{Tr}(\hat{A}) = \sum_n \langle \phi_n \mid \hat{A} \mid \phi_n \rangle = \sum_n \langle \phi'_n \mid \hat{A} \mid \phi'_n \rangle. \tag{2.285}$$

2.5.3 Matrix Representation of the Eigenvalue Problem

At issue here is to work out the matrix representation of the eigenvalue problem (2.159) and then solve it. That is, we want to find the eigenvalues a and the eigenvectors $\mid \psi \rangle$ of an operator $\hat{A}$ such that

$$\hat{A} \mid \psi \rangle = a \mid \psi \rangle, \tag{2.286}$$

where a is a complex number. Inserting the unit operator between $\hat{A}$ and $\mid \psi \rangle$ and multiplying by $\langle \phi_m \mid$, we can cast the eigenvalue equation in the form

$$\langle \phi_m \mid \hat{A} \left(\sum_n \mid \phi_n \rangle \langle \phi_n \mid \right) \mid \psi \rangle = a \langle \phi_m \mid \left(\sum_n \mid \phi_n \rangle \langle \phi_n \mid \right) \mid \psi \rangle, \tag{2.287}$$

or

$$\sum_n A_{mn} \langle \phi_n \mid \psi \rangle = a \sum_n \langle \phi_n \mid \psi \rangle \delta_{nm}, \tag{2.288}$$

which can be rewritten as

$$\sum_n [A_{mn} - a\delta_{nm}] \langle \phi_n \mid \psi \rangle = 0, \tag{2.289}$$

with $A_{mn} = \langle \phi_m \mid \hat{A} \mid \phi_n \rangle$.

This equation represents an infinite, homogeneous system of equations for the coefficients $\langle \phi_n \mid \psi \rangle$, since the basis $\{ \mid \phi_n \rangle \}$ is made of an infinite number of base kets. This system of equations can have nonzero solutions only if its determinant vanishes:

$$\det (A_{mn} - a\delta_{nm}) = 0. \tag{2.290}$$

The problem that arises here is that this determinant corresponds to a matrix with an infinite number of columns and rows. To solve (2.290) we need to truncate the basis $\{ \mid \phi_n \rangle \}$ and assume that it contains only N terms, where N must be large enough to guarantee convergence. In this case we can reduce (2.290) to the following Nth degree determinant:

$$\begin{vmatrix} A_{11} - a & A_{12} & A_{13} & \cdots & A_{1N} \\ A_{21} & A_{22} - a & A_{23} & \cdots & A_{2N} \\ A_{31} & A_{32} & A_{33} - a & \cdots & A_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & A_{N3} & \cdots & A_{NN} - a \end{vmatrix} = 0. \tag{2.291}$$

This is known as the *secular* or *characteristic equation*. The solutions of this equation yield the N *eigenvalues* $a_1, a_2, a_3, \ldots, a_N$, since it is an Nth order equation in a. The set of these N eigenvalues is called the spectrum of $\hat{A}$. Knowing the set of eigenvalues $a_1, a_2, a_3, \ldots, a_N$, we can easily determine the corresponding set of *eigenvectors* $|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_N\rangle$. For each eigenvalue a_m of $\hat{A}$, we can obtain from the "secular" equation (2.291) the N components $\langle\phi_1 | \psi\rangle, \langle\phi_2 | \psi\rangle, \langle\phi_3 | \psi\rangle, \ldots, \langle\phi_N | \psi\rangle$ of the corresponding eigenvector $|\phi_m\rangle$.

If a number of different eigenvectors (two or more) have the same eigenvalue, this eigenvalue is said to be *degenerate*. The order of degeneracy is determined by the number of linearly independent eigenvectors that have the same eigenvalue. For instance, if an eigenvalue has five different eigenvectors, it is said to be fivefold degenerate.

In the case where the set of eigenvectors $|\phi_n\rangle$ of $\hat{A}$ is complete and orthonormal, this set can be used as a basis. In this basis the matrix representing the operator $\hat{A}$ is diagonal,

$$A = \begin{pmatrix} a_1 & 0 & 0 & \cdots \\ 0 & a_2 & 0 & \cdots \\ 0 & 0 & a_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2.292}$$

the diagonal elements being the eigenvalues a_n of $\hat{A}$, since

$$\langle\phi_m | \hat{A} | \phi_n\rangle = a_n\langle\phi_m | \phi_n\rangle = a_n\delta_{mn}. \tag{2.293}$$

Note that the trace and determinant of a matrix are given, respectively, by the sum and product of the *eigenvalues*:

$$\mathrm{Tr}(A) = \sum_n a_n = a_1 + a_2 + a_3 + \cdots, \tag{2.294}$$

$$\det(A) = \prod_n a_n = a_1 a_2 a_3 \cdots. \tag{2.295}$$

Properties of determinants

Let us mention several useful properties that pertain to determinants:

$$\det(A^*) = (\det(A))^*, \qquad \det(A^\dagger) = (\det(A))^*, \tag{2.296}$$

$$\det(A^T) = \det(A), \qquad \det(A) = e^{\mathrm{Tr}(\ln A)}. \tag{2.297}$$

If A is an $N \times N$ matrix, we have

$$\det(kA) = k^N \det(A), \tag{2.298}$$

where k is a constant. In particular, if A is an $N \times N$ diagonal matrix with eigenvalues $a_1, a_2, \cdots, a_N$, it is easy to see that Equation (2.295) leads to:

$$\det(kA) = \prod_n (ka_n) = k^N \prod_n a_n = k^N a_1 a_2 a_3 \cdots a_N = k^N \det(A). \tag{2.299}$$

Another important property consists in the fact that the determinant of a product of matrices is equal to the product of their determinants:

$$\det(ABCD\cdots) = \det(A) \cdot \det(B) \cdot \det(C) \cdot \det(D) \cdots. \tag{2.300}$$

Some theorems pertaining to the eigenvalue problem
Here is a list of useful theorems (the proofs are left as exercises):

- The eigenvalues of a symmetric matrix are real; the eigenvectors form an orthonormal basis.
- The eigenvalues of an antisymmetric matrix are purely imaginary or zero.
- The eigenvalues of a Hermitian matrix are real; the eigenvectors form an orthonormal basis.
- The eigenvalues of a skew-Hermitian matrix are purely imaginary or zero.
- The eigenvalues of a unitary matrix have absolute value equal to one.
- If the eigenvalues of a square matrix are not degenerate (distinct), the corresponding eigenvectors form a basis (i.e., they form a linearly independent set).

Example 2.21 (Eigenvalues and eigenvectors of a matrix)
Find the eigenvalues and the normalized eigenvectors of the matrix

$$A = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}.$$

Solution
To find the eigenvalues of A, we simply need to solve the secular equation $\det(A - aI) = 0$:

$$0 = \begin{vmatrix} 7-a & 0 & 0 \\ 0 & 1-a & -i \\ 0 & i & -1-a \end{vmatrix} = (7-a)\left[-(1-a)(1+a) + i^2\right] = (7-a)(a^2-2). \tag{2.301}$$

The eigenvalues of A are thus given by

$$a_1 = 7, \quad a_2 = \sqrt{2}, \quad a_3 = -\sqrt{2}. \tag{2.302}$$

Let us now calculate the eigenvectors of A. To find the eigenvector corresponding to the first eigenvalue, $a_1 = 7$, we need to solve the matrix equation

$$\begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 7\begin{pmatrix} x \\ y \\ z \end{pmatrix} \Longrightarrow \begin{array}{rcl} 7x & = & 7x \\ y - iz & = & 7y \\ iy - z & = & 7z \end{array}; \tag{2.303}$$

this yields $x = 1$ (because the eigenvector is normalized) and $y = z = 0$. So the eigenvector corresponding to $a_1 = 7$ is given by the column matrix

$$|a_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{2.304}$$

This eigenvector is normalized since $\langle a_1 \mid a_1 \rangle = 1$.

The eigenvector corresponding to the second eigenvalue, $a_2 = \sqrt{2}$, can be obtained from the matrix equation

$$\begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \sqrt{2}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \Longrightarrow \begin{array}{rcl} (7-\sqrt{2})x & = & 0 \\ (1-\sqrt{2})y - iz & = & 0 \\ iy - (1+\sqrt{2})z & = & 0 \end{array}; \tag{2.305}$$

this yields $x = 0$ and $z = i(\sqrt{2} - 1)y$. So the eigenvector corresponding to $a_2 = \sqrt{2}$ is given by the column matrix

$$\mid a_2 \rangle = \begin{pmatrix} 0 \\ y \\ i(\sqrt{2}-1)y \end{pmatrix}. \tag{2.306}$$

The value of the variable y can be obtained from the normalization condition of $\mid a_2 \rangle$:

$$1 = \langle a_2 \mid a_2 \rangle = \begin{pmatrix} 0 & y^* & -i(\sqrt{2}-1)y^* \end{pmatrix}\begin{pmatrix} 0 \\ y \\ i(\sqrt{2}-1)y \end{pmatrix} = 2(2-\sqrt{2}) \mid y \mid^2 . \tag{2.307}$$

Taking only the positive value of y (a similar calculation can be performed easily if one is interested in the negative value of y), we have $y = 1/\sqrt{2(2-\sqrt{2})}$; hence the eigenvector (2.306) becomes

$$\mid a_2 \rangle = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2(2-\sqrt{2})}} \\ \frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} \end{pmatrix}. \tag{2.308}$$

Following the same procedure that led to (2.308), we can show that the third eigenvector is given by

$$\mid a_3 \rangle = \begin{pmatrix} 0 \\ y \\ -i(1+\sqrt{2})y \end{pmatrix}; \tag{2.309}$$

its normalization leads to $y = 1/\sqrt{2(2+\sqrt{2})}$ (we have considered only the positive value of y); hence

$$\mid a_3 \rangle = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2(2+\sqrt{2})}} \\ -\frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix}. \tag{2.310}$$

2.6 Representation in Continuous Bases

In this section we are going to consider the representation of state vectors, bras, and operators in *continuous* bases. After presenting the general formalism, we will consider two important applications: representations in the *position* and *momentum* spaces.

In the previous section we saw that the representations of kets, bras, and operators in a discrete basis are given by discrete matrices. We will show here that these quantities are represented in a *continuous* basis by *continuous matrices*, that is, by noncountable infinite matrices.

2.6.1 General Treatment

The orthonormality condition of the base *kets* of the continuous basis $|\chi_k\rangle$ is expressed not by the usual discrete Kronecker delta as in (2.203) but by Dirac's *continuous delta* function:

$$\langle\chi_k|\chi_{k'}\rangle = \delta(k'-k), \tag{2.311}$$

where k and k' are continuous parameters and where $\delta(k'-k)$ is the Dirac delta function (see Appendix A), which is defined by

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ikx}dk. \tag{2.312}$$

As for the completeness condition of this continuous basis, it is not given by a discrete sum as in (2.205), but by an integral over the continuous variable

$$\int_{-\infty}^{+\infty} dk\ |\chi_k\rangle\langle\chi_k| = \hat{I}, \tag{2.313}$$

where $\hat{I}$ is the unit operator.

Every state vector $|\psi\rangle$ can be expanded in terms of the complete set of basis *kets* $|\chi_k\rangle$:

$$|\psi\rangle = \hat{I}|\psi\rangle = \left(\int_{-\infty}^{+\infty} dk\ |\chi_k\rangle\langle\chi_k|\right)|\psi\rangle = \int_{-\infty}^{+\infty} dk\, b(k)\,|\chi_k\rangle, \tag{2.314}$$

where $b(k)$, which is equal to $\langle\chi_k|\psi\rangle$, represents the projection of $|\psi\rangle$ on $|\chi_k\rangle$.

The norm of the discrete base *kets* is finite ($\langle\phi_n|\phi_n\rangle = 1$), but the norm of the continuous base *kets* is infinite; a combination of (2.311) and (2.312) leads to

$$\langle\chi_k|\chi_k\rangle = \delta(0) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk \quad \longrightarrow \quad \infty. \tag{2.315}$$

This implies that the kets $|\chi_k\rangle$ are not square integrable and hence are not elements of the Hilbert space; recall that the space spanned by square-integrable functions is a Hilbert space. Despite the divergence of the norm of $|\chi_k\rangle$, the set $|\chi_k\rangle$ does constitute a valid basis of vectors that span the Hilbert space, since for any state vector $|\psi\rangle$, the scalar product $\langle\chi_k|\psi\rangle$ is finite.

The Dirac delta function

Before dealing with the representation of kets, bras, and operators, let us make a short detour to list some of the most important properties of the Dirac delta function (for a more detailed presentation, see Appendix A):

$$\delta(x) = 0, \qquad \text{for} \qquad x \neq 0, \tag{2.316}$$

$$\int_a^b f(x)\delta(x - x_0)\,dx = \begin{cases} f(x_0) & \text{if } a < x_0 < b, \\ 0 & \text{elsewhere,} \end{cases} \tag{2.317}$$

$$\int_{-\infty}^{\infty} f(x)\frac{d^n\delta(x-a)}{dx^n}\,dx = (-1)^n \left.\frac{d^n f(x)}{dx^n}\right|_{x=a}, \tag{2.318}$$

$$\delta(\vec{r} - \vec{r}') = \delta(x - x')\delta(y - y')\delta(z - z') = \frac{1}{r^2 \sin\theta}\delta(r - r')\delta(\theta - \theta')\delta(\varphi - \varphi'). \tag{2.319}$$

Representation of kets, bras, and operators

The representation of kets, bras, and operators can be easily inferred from the study that was carried out in the previous section, for the case of a discrete basis. For instance, the ket $|\,\psi\rangle$ is represented by a single column matrix which has a continuous (noncountable) and infinite number of components (rows) $b(k)$:

$$|\,\psi\rangle \longrightarrow \begin{pmatrix} \vdots \\ \langle\chi_k\,|\,\psi\rangle \\ \vdots \end{pmatrix}. \tag{2.320}$$

The bra $\langle\psi\,|$ is represented by a single row matrix which has a continuous (noncountable) and infinite number of components (columns):

$$\langle\psi\,| \longrightarrow (\cdots\cdots \quad \langle\psi\,|\,\chi_k\rangle \quad \cdots\cdots). \tag{2.321}$$

Operators are represented by square continuous matrices whose rows and columns have continuous and infinite numbers of components:

$$\hat{A} \longrightarrow \begin{pmatrix} \ddots & \vdots & \ddots \\ \cdots & A(k,k') & \cdots \\ \ddots & \vdots & \ddots \end{pmatrix}. \tag{2.322}$$

As an application, we are going to consider the representations in the position and momentum bases.

2.6.2 Position Representation

In the position representation, the basis consists of an infinite set of vectors $\{|\,\vec{r}\rangle\}$ which are eigenkets to the position operator $\hat{\vec{R}}$:

$$\hat{\vec{R}}\,|\,\vec{r}\rangle = \vec{r}\,|\,\vec{r}\rangle, \tag{2.323}$$

where $\vec{r}$ (without a hat), the position vector, is the eigenvalue of the operator $\hat{\vec{R}}$. The orthonormality and completeness conditions are respectively given by

$$\langle \vec{r} \mid \vec{r}' \rangle = \delta(\vec{r} - \vec{r}') = \delta(x - x')\delta(y - y')\delta(z - z'), \tag{2.324}$$

$$\int d^3 r \mid \vec{r} \rangle \langle \vec{r} \mid = \hat{I}, \tag{2.325}$$

since, as discussed in Appendix A, the three-dimensional delta function is given by

$$\delta(\vec{r} - \vec{r}') = \frac{1}{(2\pi)^3} \int d^3 k \, e^{i\vec{k}\cdot(\vec{r} - \vec{r}')}. \tag{2.326}$$

So every state vector $\mid \psi \rangle$ can be expanded as follows:

$$\mid \psi \rangle = \int d^3 r \mid \vec{r} \rangle \langle \vec{r} \mid \psi \rangle \equiv \int d^3 r \, \psi(\vec{r}) \mid \vec{r} \rangle, \tag{2.327}$$

where $\psi(\vec{r})$ denotes the components of $\mid \psi \rangle$ in the $\{\mid \vec{r} \rangle\}$ basis:

$$\langle \vec{r} \mid \psi \rangle = \psi(\vec{r}). \tag{2.328}$$

This is known as the *wave function* for the state vector $\mid \psi \rangle$. Recall that, according to the probabilistic interpretation of Born, the quantity $\mid \langle \vec{r} \mid \psi \rangle \mid^2 d^3 r$ represents the probability of finding the system in the volume element $d^3 r$.

The scalar product between two state vectors, $\mid \psi \rangle$ and $\mid \phi \rangle$, can be expressed in this form:

$$\langle \phi \mid \psi \rangle = \langle \phi \mid \left(\int d^3 r \mid \vec{r} \rangle \langle \vec{r} \mid \right) \mid \psi \rangle = \int d^3 r \, \phi^*(\vec{r}) \psi(\vec{r}). \tag{2.329}$$

Since $\hat{\vec{R}} \mid \vec{r} \rangle = \vec{r} \mid \vec{r} \rangle$ we have

$$\langle \vec{r}' \mid \hat{\vec{R}}^n \mid \vec{r} \rangle = \vec{r}^{\,n} \delta(\vec{r}' - \vec{r}). \tag{2.330}$$

Note that the operator $\hat{\vec{R}}$ is Hermitian, since

$$\begin{aligned} \langle \phi \mid \hat{\vec{R}} \mid \psi \rangle &= \int d^3 r \, \vec{r} \langle \phi \mid \vec{r} \rangle \langle \vec{r} \mid \psi \rangle = \left[\int d^3 r \, \vec{r} \langle \psi \mid \vec{r} \rangle \langle \vec{r} \mid \phi \rangle \right]^* \\ &= \langle \psi \mid \hat{\vec{R}} \mid \phi \rangle^*. \end{aligned} \tag{2.331}$$

2.6.3 Momentum Representation

The basis $\{\mid \vec{p} \rangle\}$ of the momentum representation is obtained from the eigenkets of the momentum operator $\hat{\vec{P}}$:

$$\hat{\vec{P}} \mid \vec{p} \rangle = \vec{p} \mid \vec{p} \rangle, \tag{2.332}$$

where $\vec{p}$ is the momentum vector. The algebra relevant to this representation can be easily inferred from the position representation. The orthonormality and completeness conditions of the momentum space basis $|\vec{p}\rangle$ are given by

$$\langle\vec{p}\,|\,\vec{p}\,'\rangle = \delta(\vec{p}-\vec{p}\,') \quad \text{and} \quad \int d^3p\ |\,\vec{p}\rangle\langle\vec{p}\,| = \hat{I}. \tag{2.333}$$

Expanding $|\psi\rangle$ in this basis, we obtain

$$|\psi\rangle = \int d^3p\ |\,\vec{p}\rangle\langle\vec{p}\,|\,\psi\rangle = \int d^3p\,\Psi(\vec{p})\,|\,\vec{p}\rangle, \tag{2.334}$$

where the expansion coefficient $\Psi(\vec{p})$ represents the *momentum space wave function*. The quantity $|\,\Psi(\vec{p})\,|^2\,d^3p$ is the probability of finding the system's momentum in the volume element d^3p located between $\vec{p}$ and $\vec{p}+d\vec{p}$.

By analogy with (2.329) the scalar product between two states is given in the momentum space by

$$\langle\phi\,|\,\psi\rangle = \langle\phi\,|\left(\int d^3p\ |\,\vec{p}\rangle\langle\vec{p}\,|\right)|\,\psi\rangle = \int d^3p\,\Phi^*(\vec{p})\Psi(\vec{p}). \tag{2.335}$$

Since $\hat{\vec{P}}\,|\,\vec{p}\rangle = \vec{p}\,|\,\vec{p}\rangle$ we have

$$\langle\vec{p}\,'\,|\,\hat{\vec{P}}^n\,|\,\vec{p}\rangle = \vec{p}^{\,n}\delta(\vec{p}\,'-\vec{p}). \tag{2.336}$$

2.6.4 Connecting the Position and Momentum Representations

Let us now study how to establish a connection between the position and the momentum representations. By analogy with the foregoing study, when changing from the $\{|\,\vec{r}\rangle\}$ basis to the $\{|\,\vec{p}\rangle\}$ basis, we encounter the *transformation* function $\langle\vec{r}\,|\,\vec{p}\rangle$.

To find the expression for the transformation function $\langle\vec{r}\,|\,\vec{p}\rangle$, let us establish a connection between the position and momentum representations of the state vector $|\,\psi\rangle$:

$$\langle\vec{r}\,|\,\psi\rangle = \langle\vec{r}\,|\left(\int d^3p\ |\,\vec{p}\rangle\langle\vec{p}\,|\right)|\,\psi\rangle = \int d^3p\ \langle\vec{r}\,|\,\vec{p}\rangle\Psi(\vec{p}); \tag{2.337}$$

that is,

$$\psi(\vec{r}) = \int d^3p\,\langle\vec{r}\,|\,\vec{p}\rangle\Psi(\vec{p}). \tag{2.338}$$

Similarly, we can write

$$\Psi(\vec{p}) = \langle\vec{p}\,|\,\psi\rangle = \langle\vec{p}\,|\int d^3r\,|\,\vec{r}\rangle\langle\vec{r}\,|\,\psi\rangle = \int d^3r\langle\vec{p}\,|\,\vec{r}\rangle\psi(\vec{r}). \tag{2.339}$$

The last two relations imply that $\psi(\vec{r})$ and $\Psi(\vec{p})$ are to be viewed as Fourier transforms of each other. In quantum mechanics the Fourier transform of a function $f(\vec{r})$ is given by

$$f(\vec{r}) = \frac{1}{(2\pi\hbar)^{3/2}}\int d^3p\,e^{i\vec{p}\cdot\vec{r}/\hbar}g(\vec{p}); \tag{2.340}$$

notice the presence of Planck's constant. Hence the function $\langle \vec{r} \mid \vec{p} \rangle$ is given by

$$\boxed{\langle \vec{r} \mid \vec{p} \rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\vec{p}\cdot\vec{r}/\hbar}.} \tag{2.341}$$

This function transforms from the momentum to the position representation. The function corresponding to the inverse transformation, $\langle \vec{p} \mid \vec{r} \rangle$, is given by

$$\boxed{\langle \vec{p} \mid \vec{r} \rangle = \langle \vec{r} \mid \vec{p} \rangle^* = \frac{1}{(2\pi\hbar)^{3/2}} e^{-i\vec{p}\cdot\vec{r}/\hbar}.} \tag{2.342}$$

The quantity $\left|\langle \vec{r} \mid \vec{p} \rangle\right|^2$ represents the probability density of finding the particle in a region around $\vec{r}$ where its momentum is equal to $\vec{p}$.

Remark
If the position wave function

$$\psi(\vec{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3p\, e^{i\vec{p}\cdot\vec{r}/\hbar} \Psi(\vec{p}) \tag{2.343}$$

is normalized (i.e., $\int d^3r\, \psi(\vec{r})\psi^*(\vec{r}) = 1$), its Fourier transform

$$\Psi(\vec{p}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3r\, e^{-i\vec{p}\cdot\vec{r}/\hbar} \psi(\vec{r}) \tag{2.344}$$

must also be normalized, since

$$\begin{aligned} \int d^3p\, \Psi^*(\vec{p})\Psi(\vec{p}) &= \int d^3p \Psi^*(\vec{p}) \left[\frac{1}{(2\pi\hbar)^{3/2}} \int d^3r\, e^{-i\vec{p}\cdot\vec{r}/\hbar} \psi(\vec{r}) \right] \\ &= \int d^3r\, \psi(\vec{r}) \left[\frac{1}{(2\pi\hbar)^{3/2}} \int d^3p\, \Psi^*(\vec{p}) e^{-i\vec{p}\cdot\vec{r}/\hbar} \right] \\ &= \int d^3r\, \psi(\vec{r})\psi^*(\vec{r}) \\ &= 1. \end{aligned} \tag{2.345}$$

This result is known as *Parseval's theorem*.

2.6.4.1 Momentum Operator in the Position Representation

To determine the form of the momentum operator $\hat{\vec{P}}$ in the position representation, let us calculate $\langle \vec{r} \mid \hat{\vec{P}} \mid \psi \rangle$:

$$\begin{aligned} \langle \vec{r} \mid \hat{\vec{P}} \mid \psi \rangle &= \int \langle \vec{r} \mid \hat{\vec{P}} \mid \vec{p} \rangle \langle \vec{p} \mid \psi \rangle d^3p = \int \vec{p} \langle \vec{r} \mid \vec{p} \rangle \langle \vec{p} \mid \psi \rangle d^3p \\ &= \frac{1}{(2\pi\hbar)^{3/2}} \int \vec{p}\, e^{i\vec{p}\cdot\vec{r}/\hbar} \Psi(\vec{p}) d^3p, \end{aligned} \tag{2.346}$$

where we have used the relation $\int \mid \vec{p}\rangle\langle\vec{p} \mid d^3p = \hat{I}$ along with Eq. (2.341). Now, since $\vec{p}\, e^{i\vec{p}\cdot\vec{r}/\hbar} = -i\hbar\vec{\nabla}e^{i\vec{p}\cdot\vec{r}/\hbar}$, and using Eq. (2.341) again, we can rewrite (2.346) as

$$\begin{aligned}\langle\vec{r} \mid \hat{\vec{P}} \mid \psi\rangle &= -i\hbar\vec{\nabla}\left(\frac{1}{(2\pi\hbar)^{3/2}}\int e^{i\vec{p}\cdot\vec{r}/\hbar}\Psi(\vec{p})d^3p\right)\\ &= -i\hbar\vec{\nabla}\left(\int\langle\vec{r} \mid \vec{p}\rangle\langle\vec{p} \mid \psi\rangle d^3p\right)\\ &= -i\hbar\vec{\nabla}\langle\vec{r} \mid \psi\rangle. \end{aligned} \tag{2.347}$$

Thus, $\hat{\vec{P}}$ is given in the position representation by

$$\boxed{\hat{\vec{P}} = -i\hbar\vec{\nabla}.} \tag{2.348}$$

Its Cartesian components are

$$\boxed{\hat{P}_x = -i\hbar\frac{\partial}{\partial x}, \qquad \hat{P}_y = -i\hbar\frac{\partial}{\partial y}, \qquad \hat{P}_z = -i\hbar\frac{\partial}{\partial z}.} \tag{2.349}$$

Note that the form of the momentum operator (2.348) can be derived by simply applying the gradient operator $\vec{\nabla}$ on a *plane* wave function $\psi(\vec{r},t) = Ae^{i(\vec{p}\cdot\vec{r}-Et)/\hbar}$:

$$-i\hbar\vec{\nabla}\psi(\vec{r},t) = \vec{p}\psi(\vec{r},t) = \hat{\vec{P}}\psi(\vec{r},t). \tag{2.350}$$

It is easy to verify that $\hat{\vec{P}}$ is Hermitian (see equation (2.416)).

Now, since $\hat{\vec{P}} = -i\hbar\vec{\nabla}$, we can write the Hamiltonian operator $\hat{H} = \hat{\vec{P}}^{\,2}/(2m) + \hat{V}$ in the position representation as follows:

$$\boxed{\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}(\vec{r}) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + \hat{V}(\vec{r}),} \tag{2.351}$$

where ∇^2 is the Laplacian operator; it is given in Cartesian coordinates by $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$.

2.6.4.2 Position Operator in the Momentum Representation

The form of the position operator $\hat{\vec{R}}$ in the momentum representation can be easily inferred from the representation of $\hat{\vec{P}}$ in the position space. In momentum space the position operator can be written as follows:

$$\hat{R}_j = i\hbar\frac{\partial}{\partial p_j} \qquad (j = x, y, z) \tag{2.352}$$

or

$$\boxed{\hat{X} = i\hbar\frac{\partial}{\partial p_x}, \qquad \hat{Y} = i\hbar\frac{\partial}{\partial p_y}, \qquad \hat{Z} = i\hbar\frac{\partial}{\partial p_z}.} \tag{2.353}$$

2.6.4.3 Important Commutation Relations

Let us now calculate the commutator $[\hat{R}_j, \hat{P}_k]$ in the position representation. As the separate actions of $\hat{X}\hat{P}_x$ and $\hat{P}_x\hat{X}$ on the wave function $\psi(\vec{r})$ are given by

$$\hat{X}\hat{P}_x\psi(\vec{r}) = -i\hbar x\frac{\partial\psi(\vec{r})}{\partial x}, \tag{2.354}$$

$$\hat{P}_x\hat{X}\psi(\vec{r}) = -i\hbar\frac{\partial}{\partial x}(x\psi(\vec{r})) = -i\hbar\psi(\vec{r}) - i\hbar x\frac{\partial\psi(\vec{r})}{\partial x}, \tag{2.355}$$

we have

$$\begin{aligned}[\hat{X}, \hat{P}_x]\psi(\vec{r}) &= \hat{X}\hat{P}_x\psi(\vec{r}) - \hat{P}_x\hat{X}\psi(\vec{r}) = -i\hbar x\frac{\partial\psi(\vec{r})}{\partial x} + i\hbar\psi(\vec{r}) + i\hbar x\frac{\partial\psi(\vec{r})}{\partial x} \\ &= i\hbar\psi(\vec{r})\end{aligned} \tag{2.356}$$

or

$$[\hat{X}, \hat{P}_x] = i\hbar. \tag{2.357}$$

Similar relations can be derived at once for the y and the z components:

$$\boxed{[\hat{X}, \hat{P}_x] = i\hbar, \qquad [\hat{Y}, \hat{P}_Y] = i\hbar, \qquad [\hat{Z}, \hat{P}_Z] = i\hbar.} \tag{2.358}$$

We can verify that

$$[\hat{X}, \hat{P}_y] = [\hat{X}, \hat{P}_z] = [\hat{Y}, \hat{P}_x] = [\hat{Y}, \hat{P}_z] = [\hat{Z}, \hat{P}_x] = [\hat{Z}, \hat{P}_y] = 0, \tag{2.359}$$

since the x, y, z degrees of freedom are independent; the previous two relations can be grouped into

$$\boxed{[\hat{R}_j, \hat{P}_k] = i\hbar\delta_{jk}, \qquad [\hat{R}_j, \hat{R}_k] = 0, \qquad [\hat{P}_j, \hat{P}_k] = 0 \qquad (j,\ k = x, y, z).} \tag{2.360}$$

These relations are often called the *canonical commutation relations*.

Now, from (2.357) we can show that (for the proof see Problem 2.9 on page 160)

$$\boxed{[\hat{X}^n, \hat{P}_x] = i\hbar n\hat{X}^{n-1}, \qquad [\hat{X}, \hat{P}_x^n] = i\hbar n\hat{P}_x^{n-1}.} \tag{2.361}$$

Following the same procedure that led to (2.356), we can obtain a more general commutation relation of $\hat{P}_x$ with an arbitrary function $f(\hat{X})$:

$$\boxed{[f(\hat{X}), \hat{P}_x] = i\hbar\frac{df(\hat{X})}{d\hat{X}} \quad \Longrightarrow \quad \left[\hat{\vec{P}},\ F(\hat{\vec{R}})\right] = -i\hbar\vec{\nabla}F(\hat{\vec{R}}),} \tag{2.362}$$

where F is a function of the operator $\hat{\vec{R}}$.

The explicit form of operators thus depends on the representation adopted. We have seen, however, that the *commutation relations for operators are representation independent*. In particular, the commutator $[\hat{R}_j, \hat{P}_k]$ is given by $i\hbar\delta_{jk}$ in the position and the momentum representations; see the next example.

Example 2.22 (Commutators are representation independent)
Calculate the commutator $[\hat{X}, \hat{P}]$ in the momentum representation and verify that it is equal to $i\hbar$.

Solution
As the operator $\hat{X}$ is given in the momentum representation by $\hat{X} = i\hbar\partial/\partial p$, we have

$$\begin{aligned}[\hat{X}, \hat{P}]\psi(p) &= \hat{X}\hat{P}\psi(p) - \hat{P}\hat{X}\psi(p) = i\hbar\frac{\partial}{\partial p}(p\psi(p)) - i\hbar p\frac{\partial\psi(p)}{\partial p} \\ &= i\hbar\psi(p) + i\hbar p\frac{\partial\psi(p)}{\partial p} - i\hbar p\frac{\partial\psi(p)}{\partial p} = i\hbar\psi(p). \end{aligned} \tag{2.363}$$

Thus, the commutator $[\hat{X}, \hat{P}]$ is given in the *momentum representation* by

$$[\hat{X}, \hat{P}] = \left[i\hbar\frac{\partial}{\partial p}, \hat{P}\right] = i\hbar. \tag{2.364}$$

The commutator $[\hat{X}, \hat{P}]$ was also shown to be equal to $i\hbar$ in the *position representation* (see equation (2.357):

$$[\hat{X}, \hat{P}] = -\left[\hat{X}, i\hbar\frac{\partial}{\partial p_x}\right] = i\hbar. \tag{2.365}$$

2.6.5 Parity Operator

The *space reflection* about the origin of the coordinate system is called an *inversion* or a *parity* operation. This transformation is *discrete*. The parity operator $\hat{\mathcal{P}}$ is defined by its action on the kets $|\vec{r}\rangle$ of the position space:

$$\hat{\mathcal{P}}|\vec{r}\rangle = |-\vec{r}\rangle, \qquad \langle\vec{r}|\hat{\mathcal{P}}^\dagger = \langle-\vec{r}|, \tag{2.366}$$

such that

$$\hat{\mathcal{P}}\psi(\vec{r}) = \psi(-\vec{r}). \tag{2.367}$$

The parity operator is Hermitian, $\hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}}$, since

$$\begin{aligned}\int d^3r\,\phi^*(\vec{r})\left[\hat{\mathcal{P}}\psi(\vec{r})\right] &= \int d^3r\phi^*(\vec{r})\psi(-\vec{r}) = \int d^3r\,\phi^*(-\vec{r})\psi(\vec{r}) \\ &= \int d^3r\,\left[\hat{\mathcal{P}}\phi(\vec{r})\right]^*\psi(\vec{r}). \end{aligned} \tag{2.368}$$

From the definition (2.367), we have

$$\hat{\mathcal{P}}^2\psi(\vec{r}) = \hat{\mathcal{P}}\psi(-\vec{r}) = \psi(\vec{r}); \tag{2.369}$$

hence $\hat{\mathcal{P}}^2$ is equal to the unit operator:

$$\hat{\mathcal{P}}^2 = \hat{I} \quad \text{or} \quad \hat{\mathcal{P}} = \hat{\mathcal{P}}^{-1}. \tag{2.370}$$

The parity operator is therefore *unitary*, since its Hermitian adjoint is equal to its inverse:

$$\hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}}^{-1}. \tag{2.371}$$

Now, since $\hat{\mathcal{P}}^2 = \hat{I}$, the eigenvalues of $\hat{\mathcal{P}}$ are $+1$ or -1 with the corresponding eigenstates

$$\hat{\mathcal{P}}\psi_+(\vec{r}) = \psi_+(-\vec{r}) = \psi_+(\vec{r}), \qquad \hat{\mathcal{P}}\psi_-(\vec{r}) = \psi_-(-\vec{r}) = -\psi_-(\vec{r}). \tag{2.372}$$

The eigenstate $|\psi_+\rangle$ is said to be *even* and $|\psi_-\rangle$ is *odd*. Therefore, the eigenfunctions of the parity operator have *definite parity*: they are either even or odd.

Since $|\psi_+\rangle$ and $|\psi_-\rangle$ are joint eigenstates of the same Hermitian operator $\hat{\mathcal{P}}$ but with different eigenvalues, these eigenstates must be orthogonal:

$$\langle\psi_+ | \psi_-\rangle = \int d^3r\, \psi_+^*(-\vec{r})\psi_-(-\vec{r}) \equiv -\int d^3r\, \psi_+^*(\vec{r})\psi_-(\vec{r}) = -\langle\psi_+ | \psi_-\rangle; \tag{2.373}$$

hence $\langle\psi_+ | \psi_-\rangle$ is zero. The states $|\psi_+\rangle$ and $|\psi_-\rangle$ form a complete set since any function can be written as $\psi(\vec{r}) = \psi_+(\vec{r}) + \psi_-(\vec{r})$, which leads to

$$\psi_+(\vec{r}) = \frac{1}{2}\left[\psi(\vec{r}) + \psi(-\vec{r})\right], \qquad \psi_-(\vec{r}) = \frac{1}{2}\left[\psi(\vec{r}) - \psi(-\vec{r})\right]. \tag{2.374}$$

Since $\hat{\mathcal{P}}^2 = I$ we have

$$\hat{\mathcal{P}}^n = \begin{cases} \hat{\mathcal{P}} & \text{when } n \text{ is odd,} \\ \hat{I} & \text{when } n \text{ is even.} \end{cases} \tag{2.375}$$

Even and odd operators

An operator $\hat{A}$ is said to be *even* if it obeys the condition

$$\hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}} = \hat{A} \tag{2.376}$$

and an operator $\hat{B}$ is *odd* if

$$\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}} = -\hat{B}. \tag{2.377}$$

We can easily verify that even operators commute with the parity operator $\hat{\mathcal{P}}$ and that odd operators anticommute with $\hat{\mathcal{P}}$:

$$\hat{A}\hat{\mathcal{P}} = (\hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}})\hat{\mathcal{P}} = \hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}}^2 = \hat{\mathcal{P}}\hat{A}, \tag{2.378}$$

$$\hat{B}\hat{\mathcal{P}} = -(\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}})\hat{\mathcal{P}} = -\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}}^2 = -\hat{\mathcal{P}}\hat{B}. \tag{2.379}$$

The fact that even operators commute with the parity operator has very useful consequences. Let us examine the following two important cases depending on whether an even operator has nondegenerate or degenerate eigenvalues:

- If an even operator is Hermitian and none of its eigenvalues is degenerate, then this operator has the same eigenvectors as those of the parity operator. And since the eigenvectors of the parity operator are either even or odd, the eigenvectors of an even, Hermitian, and nondegenerate operator must also be either even or odd; they are said to have a *definite parity*. This property will have useful applications when we solve the Schrödinger equation for even Hamiltonians.

- If the even operator has a degenerate spectrum, its eigenvectors do not necessarily have a definite parity.

What about the parity of the position and momentum operators, $\hat{\vec{R}}$ and $\hat{\vec{P}}$? We can easily show that both of them are odd, since they anticommute with the parity operator:

$$\hat{\mathcal{P}}\hat{\vec{R}} = -\hat{\vec{R}}\hat{\mathcal{P}}, \qquad \hat{\mathcal{P}}\hat{\vec{P}} = -\hat{\vec{P}}\hat{\mathcal{P}}; \tag{2.380}$$

hence

$$\hat{\mathcal{P}}\hat{\vec{R}}\hat{\mathcal{P}}^\dagger = -\hat{\vec{R}}, \qquad \hat{\mathcal{P}}\hat{\vec{P}}\hat{\mathcal{P}}^\dagger = -\hat{\vec{P}}, \tag{2.381}$$

since $\hat{\mathcal{P}}\hat{\mathcal{P}}^\dagger = 1$. For instance, to show that $\hat{\vec{R}}$ anticommutes with $\hat{\mathcal{P}}$, we need simply to look at the following relations:

$$\hat{\mathcal{P}}\hat{\vec{R}} \,|\, \vec{r}\rangle = \vec{r}\hat{\mathcal{P}} \,|\, \vec{r}\rangle = \vec{r} \,|\, -\vec{r}\rangle, \tag{2.382}$$

$$\hat{\vec{R}}\hat{\mathcal{P}} \,|\, \vec{r}\rangle = \hat{\vec{R}} \,|\, -\vec{r}\rangle = -\vec{r} \,|\, -\vec{r}\rangle. \tag{2.383}$$

If the operators $\hat{A}$ and $\hat{B}$ are even and odd, respectively, we can verify that

$$\hat{\mathcal{P}}\hat{A}^n\hat{\mathcal{P}} = \hat{A}^n, \qquad \hat{\mathcal{P}}\hat{B}^n\hat{\mathcal{P}} = (-1)^n\hat{B}^n. \tag{2.384}$$

These relations can be shown as follows:

$$\hat{\mathcal{P}}\hat{A}^n\hat{\mathcal{P}} = (\hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}})(\hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}})\cdots(\hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}}) = \hat{A}^n, \tag{2.385}$$

$$\hat{\mathcal{P}}\hat{B}^n\hat{\mathcal{P}} = (\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}})(\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}})\cdots(\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}}) = (-1)^n\hat{B}^n. \tag{2.386}$$

2.7 Matrix and Wave Mechanics

In this chapter we have so far worked out the mathematics pertaining to quantum mechanics in two different representations: *discrete* basis systems and *continuous* basis systems. The theory of quantum mechanics deals in essence with solving the following eigenvalue problem:

$$\hat{H} \,|\, \psi\rangle = E \,|\, \psi\rangle, \tag{2.387}$$

where $\hat{H}$ is the Hamiltonian of the system. This equation is general and does not depend on any coordinate system or representation. But to solve it, we need to represent it in a given basis system. The complexity associated with solving this eigenvalue equation will then vary from one basis to another.

In what follows we are going to examine the representation of this eigenvalue equation in a *discrete* basis and then in a *continuous* basis.

2.7.1 Matrix Mechanics

The representation of quantum mechanics in a *discrete* basis yields a *matrix* eigenvalue problem. That is, the representation of (2.387) in a discrete basis $\{|\phi_n\rangle\}$ yields the following matrix eigenvalue equation (see (2.291)):

$$\begin{vmatrix} H_{11}-E & H_{12} & H_{13} & \cdots & H_{1N} \\ H_{21} & H_{22}-E & H_{23} & \cdots & H_{2N} \\ H_{31} & H_{32} & H_{33}-E & \cdots & H_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{N1} & H_{N2} & H_{N3} & \cdots & H_{NN}-E \end{vmatrix} = 0. \tag{2.388}$$

This is an Nth order equation in E; its solutions yield the energy spectrum of the system: E_1, E_2, E_3, ..., E_N. Knowing the set of eigenvalues E_1, E_2, E_3, ..., E_N, we can easily determine the corresponding set of eigenvectors $|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_N\rangle$.

The diagonalization of the Hamiltonian matrix (2.388) of a system yields the energy spectrum as well as the state vectors of the system. This procedure, which was worked out by Heisenberg, involves only matrix quantities and matrix eigenvalue equations. This formulation of quantum mechanics is known as *matrix mechanics*.

The starting point of Heisenberg, in his attempt to find a theoretical foundation to Bohr's ideas, was the atomic transition relation, $\nu_{mn} = (E_m - E_n)/h$, which gives the frequencies of the radiation associated with the electron's transition from orbit m to orbit n. The frequencies ν_{mn} can be arranged in a square matrix, where the mn element corresponds to the transition from the mth to the nth quantum state.

We can also construct matrices for other dynamical quantities related to the transition $m \to n$. In this way, every physical quantity is represented by a matrix. For instance, we represent the energy levels by an energy matrix, the position by a position matrix, the momentum by a momentum matrix, the angular momentum by an angular momentum matrix, and so on. In calculating the various physical magnitudes, one has thus to deal with the algebra of matrix quantities. So, within the context of matrix mechanics, one deals with noncommuting quantities, for the product of matrices does not commute. This is an essential feature that distinguishes matrix mechanics from classical mechanics, where all the quantities commute. Take, for instance, the position and momentum quantities. While commuting in classical mechanics, $px = xp$, they do not commute within the context of matrix mechanics; they are related by the commutation relation $[\hat{X}, \hat{P}_x] = i\hbar$. The same thing applies for the components of angular momentum. We should note that the role played by the commutation relations within the context of matrix mechanics is similar to the role played by Bohr's quantization condition in atomic theory. Heisenberg's matrix mechanics therefore requires the introduction of some mathematical machinery—linear vector spaces, Hilbert space, commutator algebra, and matrix algebra—that is entirely different from the mathematical machinery of classical mechanics. Here lies the justification for having devoted a somewhat lengthy section, Section 2.5, to study the matrix representation of quantum mechanics.

2.7.2 Wave Mechanics

Representing the formalism of quantum mechanics in a *continuous* basis yields an eigenvalue problem not in the form of a matrix equation, as in Heisenberg's formulation, but in the form

of a *differential equation*. The representation of the eigenvalue equation (2.387) in the *position* space yields

$$\langle \vec{r} \mid \hat{H} \mid \psi \rangle = E\langle \vec{r} \mid \psi \rangle. \tag{2.389}$$

As shown in (2.351), the Hamiltonian is given in the position representation by $-\hbar^2\nabla^2/(2m) + \hat{V}(\vec{r})$, so we can rewrite (2.389) in a more familiar form:

$$\boxed{-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r}) + \hat{V}(\vec{r})\psi(\vec{r}) = E\psi(\vec{r}),} \tag{2.390}$$

where $\langle \vec{r} \mid \psi \rangle = \psi(\vec{r})$ is the *wave function* of the system. This differential equation is known as the *Schrödinger equation* (its origin will be discussed in Chapter 3). Its solutions yield the energy spectrum of the system as well as its wave function. This formulation of quantum mechanics in *the position representation* is called *wave mechanics*.

Unlike Heisenberg, Schödinger took an entirely different starting point in his quest to find theoretical justification to Bohr's ideas. He started from the de Broglie particle–wave hypothesis and extended it to the electrons orbiting around the nucleus. Schrödinger aimed at finding an equation that describes the motion of the electron within an atom. Here the focus is on the *wave* aspect of the electron. We can show, as we did in Chapter 1, that the Bohr quantization condition, $L = n\hbar$, is equivalent to the de Broglie relation, $\lambda = 2\pi\hbar/p$. To establish this connection, we need simply to make three assumptions: (a) the wavelength of the wave associated with the orbiting electron is connected to the electron's linear momentum p by $\lambda = 2\pi\hbar/p$, (b) the electron's orbit is circular, and (c) the circumference of the electron's orbit is an integer multiple of the electron's wavelength, i.e., $2\pi r = n\lambda$. This leads at once to $2\pi r = n\times(2\pi\hbar/p)$ or $n\hbar = rp \equiv L$. This means that, for every orbit, there is only one wavelength (or one wave) associated with the electron while revolving in that orbit. This wave can be described by means of a *wave function*. So Bohr's quantization condition implies, in essence, a uniqueness of the wave function for each orbit of the electron. In Chapter 3 we will show how Schrödinger obtained his differential equation (2.390) to describe the motion of an electron in an atom.

2.8 Concluding Remarks

Historically, the matrix formulation of quantum mechanics was worked out by Heisenberg shortly before Schrödinger introduced his wave theory. The equivalence between the matrix and wave formulations was proved a few years later by using the theory of unitary transformations. Different in form, yet identical in contents, wave mechanics and matrix mechanics achieve the same goal: finding the energy spectrum and the states of quantum systems.

The matrix formulation has the advantage of greater (formal) generality, yet it suffers from a number of disadvantages. On the conceptual side, it offers no visual idea about the structure of the atom; it is less intuitive than wave mechanics. On the technical side, it is difficult to use in some problems of relative ease such as finding the stationary states of atoms. Matrix mechanics, however, becomes powerful and practical in solving problems such as the harmonic oscillator or in treating the formalism of angular momentum.

But most of the efforts of quantum mechanics focus on solving the Schrödinger equation, not the Heisenberg matrix eigenvalue problem. So in the rest of this text we deal mostly with

wave mechanics. Matrix mechanics is used only in a few problems, such as the harmonic oscillator, where it is more suitable than Schrödinger's wave mechanics.

In wave mechanics we need only to specify the potential in which the particle moves; the Schrödinger equation takes care of the rest. That is, knowing $\hat{V}(\vec{r})$, we can in principle solve equation (2.390) to obtain the various energy levels of the particle and their corresponding wave functions. The complexity we encounter in solving the differential equation depends entirely on the form of the potential; the simpler the potential the easier the solution. Exact solutions of the Schrödinger equation are possible only for a few idealized systems; we deal with such systems in Chapters 4 and 6. However, exact solutions are generally not possible, for real systems do not yield themselves to exact solutions. In such cases one has to resort to approximate solutions. We deal with such approximate treatments in Chapters 9 and 10; Chapter 9 deals with time-independent potentials and Chapter 10 with time-dependent potentials.

Before embarking on the applications of the Schrödinger equation, we need first to lay down the theoretical foundations of quantum mechanics. We take up this task in Chapter 3, where we deal with the postulates of the theory as well as their implications; the postulates are the bedrock on which the theory is built.

2.9 Solved Problems

Problem 2.1

Consider the states $|\psi\rangle = 9i\,|\phi_1\rangle + 2\,|\phi_2\rangle$ and $|\chi\rangle = -\frac{i}{\sqrt{2}}\,|\phi_1\rangle + \frac{1}{\sqrt{2}}\,|\phi_2\rangle$, where the two vectors $|\phi_1\rangle$ and $|\phi_2\rangle$ form a complete and orthonormal basis.

(a) Calculate the operators $|\psi\rangle\langle\chi|$ and $|\chi\rangle\langle\psi|$. Are they equal?

(b) Find the Hermitian conjugates of $|\psi\rangle$, $|\chi\rangle$, $|\psi\rangle\langle\chi|$, and $|\chi\rangle\langle\psi|$.

(c) Calculate $\mathrm{Tr}(|\psi\rangle\langle\chi|)$ and $\mathrm{Tr}(|\chi\rangle\langle\psi|)$. Are they equal?

(d) Calculate $|\psi\rangle\langle\psi|$ and $|\chi\rangle\langle\chi|$ and the traces $\mathrm{Tr}(|\psi\rangle\langle\psi|)$ and $\mathrm{Tr}(|\chi\rangle\langle\chi|)$. Are they projection operators?

Solution

(a) The bras corresponding to $|\psi\rangle = 9i\,|\phi_1\rangle + 2\,|\phi_2\rangle$ and $|\chi\rangle = -i\,|\phi_1\rangle/\sqrt{2} + |\phi_2\rangle/\sqrt{2}$ are given by $\langle\psi| = -9i\langle\phi_1| + 2\langle\phi_2|$ and $\langle\chi| = \frac{i}{\sqrt{2}}\langle\phi_1| + \frac{1}{\sqrt{2}}\langle\phi_2|$, respectively. Hence we have

$$\begin{aligned} |\psi\rangle\langle\chi| &= \frac{1}{\sqrt{2}}\left(9i\,|\phi_1\rangle + 2\,|\phi_2\rangle\right)\left(i\langle\phi_1| + \langle\phi_2|\right) \\ &= \frac{1}{\sqrt{2}}\left(-9\,|\phi_1\rangle\langle\phi_1| + 9i\,|\phi_1\rangle\langle\phi_2| + 2i\,|\phi_2\rangle\langle\phi_1| + 2\,|\phi_2\rangle\langle\phi_2|\right), \end{aligned} \tag{2.391}$$

$$|\chi\rangle\langle\psi| = \frac{1}{\sqrt{2}}\left(-9\,|\phi_1\rangle\langle\phi_1| - 2i\,|\phi_1\rangle\langle\phi_2| - 9i\,|\phi_2\rangle\langle\phi_1| + 2\,|\phi_2\rangle\langle\phi_2|\right). \tag{2.392}$$

As expected, $|\psi\rangle\langle\chi|$ and $|\chi\rangle\langle\psi|$ are not equal; they would be equal only if the states $|\psi\rangle$ and $|\chi\rangle$ were proportional and the proportionality constant real.

(b) To find the Hermitian conjugates of $|\psi\rangle$, $|\chi\rangle$, $|\psi\rangle\langle\chi|$, and $|\chi\rangle\langle\psi|$, we need simply to replace the factors with their respective complex conjugates, the bras with kets, and the kets with bras:

$$|\psi\rangle^\dagger = \langle\psi| = -9i\langle\phi_1| + 2\langle\phi_2|, \qquad |\chi\rangle^\dagger = \langle\chi| = \frac{1}{\sqrt{2}}\left(i\langle\phi_1| + \langle\phi_2|\right), \tag{2.393}$$

$$\begin{aligned}(|\psi\rangle\langle\chi|)^\dagger = |\chi\rangle\langle\psi| &= \frac{1}{\sqrt{2}}\left(-9|\phi_1\rangle\langle\phi_1| - 2i|\phi_1\rangle\langle\phi_2|\right. \\ &\quad \left. - 9i|\phi_2\rangle\langle\phi_1| + 2|\phi_2\rangle\langle\phi_2|\right), \end{aligned} \tag{2.394}$$

$$\begin{aligned}(|\chi\rangle\langle\psi|)^\dagger = |\psi\rangle\langle\chi| &= \frac{1}{\sqrt{2}}\left(-9|\phi_1\rangle\langle\phi_1| + 9i|\phi_1\rangle\langle\phi_2|\right. \\ &\quad \left. + 2i|\phi_2\rangle\langle\phi_1| + 2|\phi_2\rangle\langle\phi_2|\right). \end{aligned} \tag{2.395}$$

(c) Using the property $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ and since $\langle\phi_1|\phi_1\rangle = \langle\phi_2|\phi_2\rangle = 1$ and $\langle\phi_1|\phi_2\rangle = \langle\phi_2|\phi_1\rangle = 0$, we obtain

$$\begin{aligned}\mathrm{Tr}(|\psi\rangle\langle\chi|) &= \mathrm{Tr}(\langle\chi|\psi\rangle) = \langle\chi|\psi\rangle \\ &= \left(\frac{i}{\sqrt{2}}\langle\phi_1| + \frac{1}{\sqrt{2}}\langle\phi_2|\right)\left(9i|\phi_1\rangle + 2|\phi_2\rangle\right) = -\frac{7}{\sqrt{2}}, \end{aligned} \tag{2.396}$$

$$\begin{aligned}\mathrm{Tr}(|\chi\rangle\langle\psi|) &= \mathrm{Tr}(\langle\psi|\chi\rangle) = \langle\psi|\chi\rangle \\ &= \left(-9i\langle\phi_1| + 2\langle\phi_2|\right)\left(-\frac{i}{\sqrt{2}}|\phi_1\rangle + \frac{1}{\sqrt{2}}|\phi_2\rangle\right) = -\frac{7}{\sqrt{2}} \\ &= \mathrm{Tr}(|\psi\rangle\langle\chi|). \end{aligned} \tag{2.397}$$

The traces $\mathrm{Tr}(|\psi\rangle\langle\chi|)$ and $\mathrm{Tr}(|\chi\rangle\langle\psi|)$ are equal only because the scalar product of $|\psi\rangle$ and $|\chi\rangle$ is a real number. Were this product a complex number, the traces would be different; in fact, they would be the complex conjugate of one another.

(d) The expressions $|\psi\rangle\langle\psi|$ and $|\chi\rangle\langle\chi|$ are

$$\begin{aligned}|\psi\rangle\langle\psi| &= \left(9i|\phi_1\rangle + 2|\phi_2\rangle\right)\left(-9i\langle\phi_1| + 2\langle\phi_2|\right) \\ &= 81|\phi_1\rangle\langle\phi_1| + 18i|\phi_1\rangle\langle\phi_2| - 18i|\phi_2\rangle\langle\phi_1| + 4|\phi_2\rangle\langle\phi_2|, \end{aligned} \tag{2.398}$$

$$\begin{aligned}|\chi\rangle\langle\chi| &= \frac{1}{2}\left(|\phi_1\rangle\langle\phi_1| - i|\phi_1\rangle\langle\phi_2| + i|\phi_2\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2|\right) \\ &= \frac{1}{2}\left(1 - i|\phi_1\rangle\langle\phi_2| + i|\phi_2\rangle\langle\phi_1|\right). \end{aligned} \tag{2.399}$$

In deriving (2.399) we have used the fact that the basis is complete, $|\phi_1\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2| = 1$. The traces $\mathrm{Tr}(|\psi\rangle\langle\psi|)$ and $\mathrm{Tr}(|\chi\rangle\langle\chi|)$ can then be calculated immediately:

$$\mathrm{Tr}(|\psi\rangle\langle\psi|) = \langle\psi|\psi\rangle = \left(-9i\langle\phi_1| + 2\langle\phi_2|\right)\left(9i|\phi_1\rangle + 2|\phi_2\rangle\right) = 85, \tag{2.400}$$

$$\mathrm{Tr}(|\chi\rangle\langle\chi|) = \langle\chi|\chi\rangle = \frac{1}{2}\left(i\langle\phi_1| + \langle\phi_2|\right)\left(-i|\phi_1\rangle + |\phi_2\rangle\right) = 1. \tag{2.401}$$

So $|\chi\rangle$ is normalized but $|\psi\rangle$ is not. Since $|\chi\rangle$ is normalized, we can easily ascertain that $|\chi\rangle\langle\chi|$ is a projection operator, because it is Hermitian, $(|\chi\rangle\langle\chi|)^\dagger = |\chi\rangle\langle\chi|$, and equal to its own square:

$$(|\chi\rangle\langle\chi|)^2 = |\chi\rangle\langle\chi|\chi\rangle\langle\chi| = (\langle\chi|\chi\rangle)\,|\chi\rangle\langle\chi| = |\chi\rangle\langle\chi| \,. \tag{2.402}$$

As for $|\psi\rangle\langle\psi|$, although it is Hermitian, it cannot be a projection operator since $|\psi\rangle$ is not normalized. That is, $|\psi\rangle\langle\psi|$ is not equal to its own square:

$$(|\psi\rangle\langle\psi|)^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = (\langle\psi|\psi\rangle)\,|\psi\rangle\langle\psi| = 85\,|\psi\rangle\langle\psi| \,. \tag{2.403}$$

Problem 2.2

(a) If $|\psi\rangle$ is normalized, show that it satisfies the following relation:

$$(|\psi\rangle\langle\psi|)^n = |\psi\rangle\langle\psi| \,.$$

(b) Show the following relation

$$\left(|\psi\rangle\langle\phi|\right)^n = \langle\phi|\psi\rangle^{n-1}\,|\psi\rangle\langle\phi| \,.$$

Solution

(a) Let us write $(|\psi\rangle\langle\psi|)^n$ as the product of n terms:

$$\begin{aligned}
(|\psi\rangle\langle\psi|)^n &= (|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|)\cdots(|\psi\rangle\langle\psi|) \\
&= |\psi\rangle(\langle\psi|\psi\rangle)(\langle\psi|\psi\rangle)\cdots(\langle\psi|\psi\rangle)\langle\psi| \\
&= |\psi\rangle(\langle\psi|\psi\rangle)^{n-1}\langle\psi| \\
&= |\psi\rangle\langle\psi|,
\end{aligned} \tag{2.404}$$

since $\langle\psi|\psi\rangle = 1$ due to the fact that $|\psi\rangle$ is normalized.

(b) Let us write $\left(|\psi\rangle\langle\phi|\right)^n$ as the product of n terms:

$$\begin{aligned}
\left(|\psi\rangle\langle\phi|\right)^n &= \left(|\psi\rangle\langle\phi|\right)\left(|\psi\rangle\langle\phi|\right)\left(|\psi\rangle\langle\phi|\right)\cdots\left(|\psi\rangle\langle\phi|\right) \\
&= |\psi\rangle\left(\langle\phi|\psi\rangle\right)\left(\langle\phi|\psi\rangle\right)\cdots\left(\langle\phi|\psi\rangle\right)\langle\phi| \\
&= |\psi\rangle\left(\langle\phi|\psi\rangle\right)^{n-1}\langle\phi| \\
&= \langle\phi|\psi\rangle^{n-1}\,|\psi\rangle\langle\phi|,
\end{aligned} \tag{2.405}$$

since $\langle\phi|\psi\rangle$ is a scalar; hence, it can be pulled to the far left in the last term of the equation.

Problem 2.3

(a) Find a complete and orthonormal basis for a space of the trigonometric functions of the form $\psi(\theta) = \sum_{n=0}^{N} a_n \cos(n\theta)$.

(b) Illustrate the results derived in (a) for the case $N = 5$; find the basis vectors.

Solution

(a) Since $\cos(n\theta) = \frac{1}{2}\left(e^{in\theta} + e^{-in\theta}\right)$, we can write $\sum_{n=0}^{N} a_n \cos(n\theta)$ as

$$\frac{1}{2}\sum_{n=0}^{N} a_n\left(e^{in\theta} + e^{-in\theta}\right) = \frac{1}{2}\left[\sum_{n=0}^{N} a_n e^{in\theta} + \sum_{n=-N}^{0} a_{-n} e^{in\theta}\right] = \sum_{n=-N}^{N} C_n e^{in\theta}, \tag{2.406}$$

where $C_n = a_n/2$ for $n > 0$, $C_n = a_{-n}/2$ for $n < 0$, and $C_0 = a_0$. Since any trigonometric function of the form $\psi(x) = \sum_{n=0}^{N} a_n \cos(n\theta)$ can be expressed in terms of the functions $\phi_n(\theta) = e^{in\theta}/\sqrt{2\pi}$, we can try to take the set $\phi_n(\theta)$ as a basis. As this set is complete, let us see if it is orthonormal. The various functions $\phi_n(\theta)$ are indeed orthonormal, since their scalar products are given by

$$\langle \phi_m \mid \phi_n \rangle = \int_{-\pi}^{\pi} \phi_m^*(\theta)\phi_n(\theta)d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)\theta}\, d\theta = \delta_{nm}. \tag{2.407}$$

In deriving this result, we have considered two cases: $n = m$ and $n \neq m$. First, the case $n = m$ is obvious, since $\langle \phi_n \mid \phi_n \rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta = 1$. On the other hand, when $n \neq m$ we have

$$\langle \phi_m \mid \phi_n \rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)\theta} d\theta = \frac{1}{2\pi}\frac{e^{i(n-m)\pi} - e^{-i(n-m)\pi}}{i(n-m)} = \frac{2i\sin((n-m)\pi)}{2i\pi(n-m)} = 0, \tag{2.408}$$

since $\sin((n-m)\pi) = 0$. So the functions $\phi_n(\theta) = e^{in\theta}/\sqrt{2\pi}$ form a complete and orthonormal basis. From (2.406) we see that the basis has $2N + 1$ functions $\phi_n(\theta)$; hence the dimension of this space of functions is equal to $2N + 1$.

(b) In the case where $N = 5$, the dimension of the space is equal to 11, for the basis has 11 vectors: $\phi_{-5}(\theta) = e^{-5i\theta}/\sqrt{2\pi}$, $\phi_{-4}(\theta) = e^{-4i\theta}/\sqrt{2\pi}$, ..., $\phi_0(\theta) = 1/\sqrt{2\pi}$, ..., $\phi_4(\theta) = e^{4i\theta}/\sqrt{2\pi}$, $\phi_5(\theta) = e^{5i\theta}/\sqrt{2\pi}$.

Problem 2.4

(a) Show that the sum of two projection operators cannot be a projection operator unless their product is zero.

(b) Show that the product of two projection operators cannot be a projection operator unless they commute.

Solution

Recall that an operator $\hat{P}$ is a projection operator if it satisfies $\hat{P}^\dagger = \hat{P}$ and $\hat{P}^2 = \hat{P}$.

(a) If two operators $\hat{A}$ and $\hat{B}$ are projection operators and if $\hat{A}\hat{B} = \hat{B}\hat{A}$, we want to show that $(\hat{A} + \hat{B})^\dagger = \hat{A} + \hat{B}$ and that $(\hat{A} + \hat{B})^2 = \hat{A} + \hat{B}$. First, the hermiticity is easy to ascertain since $\hat{A}$ and $\hat{B}$ are both Hermitian: $(\hat{A} + \hat{B})^\dagger = \hat{A} + \hat{B}$. Let us now look at the square of $(\hat{A} + \hat{B})$; since $\hat{A}^2 = \hat{A}$ and $\hat{B}^2 = \hat{B}$, we can write

$$(\hat{A} + \hat{B})^2 = \hat{A}^2 + \hat{B}^2 + (\hat{A}\hat{B} + \hat{B}\hat{A}) = \hat{A} + \hat{B} + (\hat{A}\hat{B} + \hat{B}\hat{A}). \tag{2.409}$$

Clearly, only when the product of $\hat{A}$ and $\hat{B}$ is zero will their sum be a projection operator.

(b) At issue here is to show that if two operators $\hat{A}$ and $\hat{B}$ are projection operators and if they commute, $[\hat{A}, \hat{B}] = 0$, their product is a projection operator. That is, we need to show that

$(\hat{A}\hat{B})^\dagger = \hat{A}\hat{B}$ and $(\hat{A}\hat{B})^2 = \hat{A}\hat{B}$. Again, since $\hat{A}$ and $\hat{B}$ are Hermitian and since they commute, we see that $(\hat{A}\hat{B})^\dagger = \hat{B}\hat{A} = \hat{A}\hat{B}$. As for the square of $\hat{A}\hat{B}$, we have

$$(\hat{A}\hat{B})^2 = (\hat{A}\hat{B})(\hat{A}\hat{B}) = \hat{A}(\hat{B}\hat{A})\hat{B} = \hat{A}(\hat{A}\hat{B})\hat{B} = \hat{A}^2\hat{B}^2 = \hat{A}\hat{B}, \tag{2.410}$$

hence the product $\hat{A}\hat{B}$ is a projection operator.

Problem 2.5
Consider a state $|\psi\rangle = \frac{1}{\sqrt{2}}|\phi_1\rangle + \frac{1}{\sqrt{5}}|\phi_2\rangle + \frac{1}{\sqrt{10}}|\phi_3\rangle$ which is given in terms of three orthonormal eigenstates $|\phi_1\rangle$, $|\phi_2\rangle$ and $|\phi_3\rangle$ of an operator $\hat{B}$ such that $\hat{B}|\phi_n\rangle = n^2|\phi_n\rangle$. Find the expectation value of $\hat{B}$ for the state $|\psi\rangle$.

Solution
Using Eq (2.68), we can write the expectation value of $\hat{B}$ for the state $|\psi\rangle$ as $\langle\hat{B}\rangle = \langle\psi \mid \hat{B} \mid \psi\rangle/\langle\psi \mid \psi\rangle$ where

$$\begin{aligned}\langle\psi \mid \psi\rangle &= \left(\frac{1}{\sqrt{2}}\langle\phi_1| + \frac{1}{\sqrt{5}}\langle\phi_2| + \frac{1}{\sqrt{10}}\langle\phi_3|\right)\left(\frac{1}{\sqrt{2}}|\phi_1\rangle + \frac{1}{\sqrt{5}}|\phi_2\rangle + \frac{1}{\sqrt{10}}|\phi_3\rangle\right)\\ &= \frac{8}{10}\end{aligned} \tag{2.411}$$

and

$$\begin{aligned}\langle\psi \mid \hat{B} \mid \psi\rangle &= \left(\frac{1}{\sqrt{2}}\langle\phi_1| + \frac{1}{\sqrt{5}}\langle\phi_2| + \frac{1}{\sqrt{10}}\langle\phi_3|\right)\hat{B}\left(\frac{1}{\sqrt{2}}|\phi_1\rangle + \frac{1}{\sqrt{5}}|\phi_2\rangle + \frac{1}{\sqrt{10}}|\phi_3\rangle\right)\\ &= \frac{1}{2} + \frac{2^2}{5} + \frac{3^2}{10}\\ &= \frac{22}{10}.\end{aligned} \tag{2.412}$$

Hence, the expectation value of $\hat{B}$ is given by

$$\langle\hat{B}\rangle = \frac{\langle\psi \mid \hat{B} \mid \psi\rangle}{\langle\psi \mid \psi\rangle} = \frac{22/10}{8/10} = \frac{11}{4}. \tag{2.413}$$

Problem 2.6

(a) Study the hermiticity of these operators: $\hat{X}$, d/dx, and id/dx. What about the complex conjugate of these operators? Are the Hermitian conjugates of the position and momentum operators equal to their complex conjugates?

(b) Use the results of (a) to discuss the hermiticity of the operators $e^{\hat{X}}$, $e^{d/dx}$, and $e^{id/dx}$.

(c) Find the Hermitian conjugate of the operator $\hat{X}d/dx$.

(d) Use the results of (a) to discuss the hermiticity of the components of the angular momentum operator (Chapter 5): $\hat{L}_x = -i\hbar\left(\hat{Y}\partial/\partial z - \hat{Z}\partial/\partial y\right)$, $\hat{L}_y = -i\hbar\left(\hat{Z}\partial/\partial x - \hat{X}\partial/\partial z\right)$, $\hat{L}_z = -i\hbar\left(\hat{X}\partial/\partial y - \hat{Y}\partial/\partial x\right)$.

Solution

(a) Using (2.80) and (2.81), and using the fact that the eigenvalues of $\hat{X}$ are real (i.e., $\hat{X}^* = \hat{X}$ or $x^* = x$), we can verify that $\hat{X}$ is Hermitian (i.e., $\hat{X}^\dagger = \hat{X}$) since

$$\begin{aligned}\langle\psi \mid \hat{X}\psi\rangle &= \int_{-\infty}^{+\infty} \psi^*(x)\,(x\psi(x))\,dx = \int_{-\infty}^{+\infty} (x\psi(x)^*)\,\psi(x)\,dx \\ &= \int_{-\infty}^{+\infty} (x\psi(x))^*\,\psi(x)\,dx = \langle \hat{X}\psi \mid \psi\rangle. \end{aligned} \tag{2.414}$$

Now, since $\psi(x)$ vanishes as $x \to \pm\infty$, an integration by parts leads to

$$\begin{aligned}\langle\psi \mid \frac{d}{dx}\psi\rangle &= \int_{-\infty}^{+\infty} \psi^*(x)\left(\frac{d\psi(x)}{dx}\right)dx = \left.\psi^*(x)\psi(x)\right|_{x=-\infty}^{x=+\infty} - \int_{-\infty}^{+\infty}\left(\frac{d\psi^*(x)}{dx}\right)\psi(x)\,dx \\ &= -\int_{-\infty}^{+\infty}\left(\frac{d\psi(x)}{dx}\right)^*\psi(x)\,dx = -\langle\frac{d}{dx}\psi \mid \psi\rangle. \end{aligned} \tag{2.415}$$

So, d/dx is anti-Hermitian: $(d/dx)^\dagger = -d/dx$. Since d/dx is anti-Hermitian, id/dx must be Hermitian, since $(id/dx)^\dagger = -i(-d/dx) = id/dx$. The results derived above are

$$\boxed{\hat{X}^\dagger = \hat{X}, \qquad \left(\frac{d}{dx}\right)^\dagger = -\frac{d}{dx}, \qquad \left(i\frac{d}{dx}\right)^\dagger = i\frac{d}{dx}.} \tag{2.416}$$

From this relation we see that the momentum operator $\hat{P} = -i\hbar d/dx$ is Hermitian: $\hat{P}^\dagger = \hat{P}$. We can also infer that, although the momentum operator is Hermitian, its complex conjugate is not equal to $\hat{P}$, since $\hat{P}^* = (-i\hbar d/dx)^* = i\hbar d/dx = -\hat{P}$. We may group these results into the following relation:

$$\boxed{\hat{X}^\dagger = \hat{X}, \quad \hat{X}^* = \hat{X}, \qquad \hat{P}^\dagger = \hat{P}, \quad \hat{P}^* = -\hat{P}.} \tag{2.417}$$

(b) Using the relations $(e^{\hat{A}})^\dagger = e^{\hat{A}^\dagger}$ and $(e^{i\hat{A}})^\dagger = e^{-i\hat{A}^\dagger}$ derived in (2.126), we infer

$$\boxed{(e^{\hat{X}})^\dagger = e^{\hat{X}}, \qquad (e^{d/dx})^\dagger = e^{-d/dx}, \qquad (e^{id/dx})^\dagger = e^{id/dx}.} \tag{2.418}$$

(c) Since $\hat{X}$ is Hermitian and d/dx is anti-Hermitian, we have

$$\left(\hat{X}\frac{d}{dx}\right)^\dagger = \left(\frac{d}{dx}\right)^\dagger(\hat{X})^\dagger = -\frac{d}{dx}\hat{X}, \tag{2.419}$$

where $d\hat{X}/dx$ is given by

$$\frac{d}{dx}\left(\hat{X}\psi(x)\right) = \left(1 + x\frac{d}{dx}\right)\psi(x); \tag{2.420}$$

hence

$$\left(\hat{X}\frac{d}{dx}\right)^\dagger = -\hat{X}\frac{d}{dx} - 1. \tag{2.421}$$

(d) From the results derived in (a), we infer that the operators $\hat{Y}$, $\hat{Z}$, $i\partial/\partial x$, and $i\partial/\partial y$ are Hermitian. We can verify that $\hat{L}_x$ is also Hermitian:

$$\hat{L}_x^\dagger = -i\hbar\left(\frac{\partial}{\partial z}\hat{Y} - \frac{\partial}{\partial y}\hat{Z}\right) = -i\hbar\left(\hat{Y}\frac{\partial}{\partial z} - \hat{Z}\frac{\partial}{\partial y}\right) = \hat{L}_x; \tag{2.422}$$

in deriving this relation, we used the fact that the y and z degrees of freedom commute (i.e., $\partial\hat{Y}/\partial z = \hat{Y}\partial/\partial z$ and $\partial\hat{Z}/\partial y = \hat{Z}\partial/\partial y$), for they are independent. Similarly, the hermiticity of $\hat{L}_y = -i\hbar\left(\hat{Z}\partial/\partial x - \hat{X}\partial/\partial z\right)$ and $\hat{L}_z = -i\hbar\left(\hat{X}\partial/\partial y - \hat{Y}\partial/\partial x\right)$ is obvious.

Problem 2.7

(a) Show that the operator $\hat{A} = i(\hat{X}^2 + 1)d/dx + i\hat{X}$ is Hermitian.

(b) Find the state $\psi(x)$ for which $\hat{A}\psi(x) = 0$ and normalize it.

(c) Calculate the probability of finding the particle (represented by $\psi(x)$) in the region: $-1 \le x \le 1$.

Solution

(a) From the previous problem we know that $\hat{X}^\dagger = \hat{X}$ and $(d/dx)^\dagger = -d/dx$. We can thus infer the Hermitian conjugate of $\hat{A}$:

$$\begin{aligned}\hat{A}^\dagger &= -i\left(\frac{d}{dx}\right)^\dagger(\hat{X}^2)^\dagger - i\left(\frac{d}{dx}\right)^\dagger - i\hat{X}^\dagger = i\left(\frac{d}{dx}\right)(\hat{X}^2) + i\left(\frac{d}{dx}\right) - i\hat{X} \\ &= i\hat{X}^2\frac{d}{dx} + i\left[\frac{d}{dx}, \hat{X}^2\right] + i\frac{d}{dx} - i\hat{X}.\end{aligned} \tag{2.423}$$

Using the relation $[\hat{B}, \hat{C}^2] = \hat{C}[\hat{B}, \hat{C}] + [\hat{B}, \hat{C}]\hat{C}$ along with $[d/dx, \hat{X}] = 1$, we can easily evaluate the commutator $[d/dx, \hat{X}^2]$:

$$\left[\frac{d}{dx}, \hat{X}^2\right] = \hat{X}\left[\frac{d}{dx}, \hat{X}\right] + \left[\frac{d}{dx}, \hat{X}\right]\hat{X} = 2\hat{X}. \tag{2.424}$$

A combination of (2.423) and (2.424) shows that $\hat{A}$ is Hermitian:

$$\hat{A}^\dagger = i(\hat{X}^2 + 1)\frac{d}{dx} + i\hat{X} = \hat{A}. \tag{2.425}$$

(b) The state $\psi(x)$ for which $\hat{A}\psi(x) = 0$, i.e.,

$$i(\hat{X}^2 + 1)\frac{d\psi(x)}{dx} + i\hat{X}\psi(x) = 0, \tag{2.426}$$

corresponds to

$$\frac{d\psi(x)}{dx} = -\frac{x}{x^2 + 1}\psi(x). \tag{2.427}$$

The solution to this equation is given by

$$\psi(x) = \frac{B}{\sqrt{x^2 + 1}}. \tag{2.428}$$

Since $\int_{-\infty}^{+\infty} dx/(x^2+1) = \pi$ we have

$$1 = \int_{-\infty}^{+\infty} |\psi(x)|^2 dx = B^2 \int_{-\infty}^{+\infty} \frac{dx}{x^2+1} = B^2\pi, \tag{2.429}$$

which leads to $B = 1/\sqrt{\pi}$ and hence $\psi(x) = \frac{1}{\sqrt{\pi(x^2+1)}}$.

(c) Using the integral $\int_{-1}^{+1} dx/(x^2+1) = \pi/2$, we can obtain the probability immediately:

$$P = \int_{-1}^{+1} |\psi(x)|^2 dx = \frac{1}{\pi} \int_{-1}^{+1} \frac{dx}{x^2+1} = \frac{1}{2}. \tag{2.430}$$

Problem 2.8
Discuss the conditions for these operators to be unitary: (a) $(1 + i\hat{A})/(1 - i\hat{A})$, (b) $(\hat{A} + i\hat{B})/\sqrt{\hat{A}^2 + \hat{B}^2}$.

Solution
An operator $\hat{U}$ is unitary if $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}$ (see (2.135)).

(a) Since

$$\left(\frac{1+i\hat{A}}{1-i\hat{A}}\right)^\dagger = \frac{1-i\hat{A}^\dagger}{1+i\hat{A}^\dagger}, \tag{2.431}$$

we see that if $\hat{A}$ is Hermitian, the expression $(1 + i\hat{A})/(1 - i\hat{A})$ is unitary:

$$\left(\frac{1+i\hat{A}}{1-i\hat{A}}\right)^\dagger \frac{1+i\hat{A}}{1-i\hat{A}} = \frac{1-i\hat{A}}{1+i\hat{A}} \frac{1+i\hat{A}}{1-i\hat{A}} = \hat{I}. \tag{2.432}$$

(b) Similarly, if $\hat{A}$ and $\hat{B}$ are Hermitian and commute, the expression $(\hat{A} + i\hat{B})/\sqrt{\hat{A}^2 + \hat{B}^2}$ is unitary:

$$\begin{aligned}\left(\frac{\hat{A}+i\hat{B}}{\sqrt{\hat{A}^2+\hat{B}^2}}\right)^\dagger \frac{\hat{A}+i\hat{B}}{\sqrt{\hat{A}^2+\hat{B}^2}} &= \frac{\hat{A}-i\hat{B}}{\sqrt{\hat{A}^2+\hat{B}^2}} \frac{\hat{A}+i\hat{B}}{\sqrt{\hat{A}^2+\hat{B}^2}} = \frac{\hat{A}^2+\hat{B}^2+i(\hat{A}\hat{B}-\hat{B}\hat{A})}{\hat{A}^2+\hat{B}^2} \\ &= \frac{\hat{A}^2+\hat{B}^2}{\hat{A}^2+\hat{B}^2} = \hat{I}.\end{aligned} \tag{2.433}$$

Problem 2.9
(a) Using the commutator $[\hat{X}, \hat{p}] = i\hbar$, show that $[\hat{X}^m, \hat{P}] = im\hbar\hat{X}^{m-1}$, with $m > 1$. Can you think of a direct way to get to the same result?

(b) Use the result of (a) to show the general relation $[F(\hat{X}), \hat{P}] = i\hbar dF(\hat{X})/d\hat{X}$, where $F(\hat{X})$ is a differentiable operator function of $\hat{X}$.

Solution

(a) We offer here three methods:

First method: Proof by induction: Assuming that $[\hat{X}^m, \hat{P}] = im\hbar\hat{X}^{m-1}$ is valid for $m = k$ (note that it holds for $n = 1$; i.e., $[\hat{X}, \hat{P}] = i\hbar$),

$$[\hat{X}^k, \hat{P}] = ik\hbar\hat{X}^{k-1}, \tag{2.434}$$

let us show that it holds for $m = k + 1$:

$$[\hat{X}^{k+1}, \hat{P}] = [\hat{X}^k\hat{X}, \hat{P}] = \hat{X}^k[\hat{X}, \hat{P}] + [\hat{X}^k, \hat{P}]\hat{X}, \tag{2.435}$$

where we have used the relation $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}$. Now, since $[\hat{X}, \hat{P}] = i\hbar$ and $[\hat{X}^k, \hat{P}] = ik\hbar\hat{X}^{k-1}$, we rewrite (2.435) as

$$[\hat{X}^{k+1}, \hat{P}] = i\hbar\hat{X}^k + (ik\hbar\hat{X}^{k-1})\hat{X} = i\hbar(k+1)\hat{X}^k. \tag{2.436}$$

So this relation is valid for any value of k, notably for $k = m - 1$:

$$\boxed{[\hat{X}^m, \hat{P}] = im\hbar\hat{X}^{m-1}.} \tag{2.437}$$

Second method: In fact, it is easy to arrive at this result *directly* through brute force as follows. Using the relation $[\hat{A}^n, \hat{B}] = \hat{A}^{n-1}[\hat{A}, \hat{B}] + [\hat{A}^{n-1}, \hat{B}]\hat{A}$ along with $[\hat{X}, \hat{P}_x] = i\hbar$, we can obtain

$$[\hat{X}^2, \hat{P}_x] = \hat{X}[\hat{X}, \hat{P}_x] + [\hat{X}, \hat{P}_x]\hat{X} = 2i\hbar\hat{X}, \tag{2.438}$$

which leads to

$$[\hat{X}^3, \hat{P}_x] = \hat{X}^2[\hat{X}, \hat{P}_x] + [\hat{X}^2, \hat{P}_x]\hat{X} = 3i\hat{X}^2\hbar; \tag{2.439}$$

this in turn leads to

$$[\hat{X}^4, \hat{P}_x] = \hat{X}^3[\hat{X}, \hat{P}_x] + [\hat{X}^3, \hat{P}_x]\hat{X} = 4i\hat{X}^3\hbar. \tag{2.440}$$

Continuing in this way, we can get to any power of $\hat{X}$: $[\hat{X}^m, \hat{P}] = im\hbar\hat{X}^{m-1}$.

Third method: A more direct and simpler method is to apply the commutator $[\hat{X}^m, \hat{P}]$ on some wave function $\psi(x)$:

$$\begin{aligned}
[\hat{X}^m, \hat{P}_x]\psi(x) &= \left(\hat{X}^m\hat{P}_x - \hat{P}_x\hat{X}^m\right)\psi(x) \\
&= x^m\left(-i\hbar\frac{d\psi(x)}{dx}\right) + i\hbar\frac{d}{dx}\left(x^m\psi(x)\right) \\
&= x^m\left(-i\hbar\frac{d\psi(x)}{dx}\right) + im\hbar x^{m-1}\psi(x) - x^m\left(-i\hbar\frac{d\psi(x)}{dx}\right) \\
&= im\hbar x^{m-1}\psi(x).
\end{aligned} \tag{2.441}$$

Since $[\hat{X}^m, \hat{P}_x]\psi(x) = im\hbar x^{m-1}\psi(x)$ we see that $[\hat{X}^m, \hat{P}] = im\hbar\hat{X}^{m-1}$.

(b) Let us Taylor expand $F(\hat{X})$ in powers of $\hat{X}$, $F(\hat{X}) = \sum_k a_k\hat{X}^k$, and insert this expression into $[F(\hat{X}), \hat{P}]$:

$$\left[F(\hat{X}), \hat{P}\right] = \left[\sum_k a_k\hat{X}^k, \hat{P}\right] = \sum_k a_k[\hat{X}^k, \hat{P}], \tag{2.442}$$

where the commutator $[\hat{X}^k, \hat{P}]$ is given by (2.434). Thus, we have

$$\left[F(\hat{X}), \hat{P}\right] = i\hbar \sum_k k a_k \hat{X}^{k-1} = i\hbar \frac{d(\sum_k a_k \hat{X}^k)}{d\hat{X}} = i\hbar \frac{dF(\hat{X})}{d\hat{X}}. \tag{2.443}$$

A much simpler method again consists in applying the commutator $\left[F(\hat{X}), \hat{P}\right]$ on some wave function $\psi(x)$. Since $F(\hat{X})\psi(x) = F(x)\psi(x)$, we have

$$\begin{aligned}
\left[F(\hat{X}), \hat{P}\right]\psi(x) &= F(\hat{X})\hat{P}\psi(x) + i\hbar \frac{d}{dx}\left(F(x)\psi(x)\right) \\
&= F(\hat{X})\hat{P}\psi(x) - \left(-i\hbar \frac{d\psi(x)}{dx}\right)F(x) + i\hbar \frac{dF(x)}{dx}\psi(x) \\
&= F(\hat{X})\hat{P}\psi(x) - F(\hat{X})\hat{P}\psi(x) + i\hbar \frac{dF(x)}{dx}\psi(x) \\
&= i\hbar \frac{dF(x)}{dx}\psi(x).
\end{aligned} \tag{2.444}$$

Since $\left[F(\hat{X}), \hat{P}\right]\psi(x) = i\hbar \frac{dF(x)}{dx}\psi(x)$ we see that $\left[F(\hat{X}), \hat{P}\right] = i\hbar \frac{dF(\hat{X})}{d\hat{X}}$.

Problem 2.10

Consider the matrices $A = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2i & 0 \\ i & 0 & -5i \end{pmatrix}$.

(a) Are A and B Hermitian? Calculate AB and BA and verify that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$; then calculate $[A, B]$ and verify that $\mathrm{Tr}([A, B]) = 0$.

(b) Find the eigenvalues and the normalized eigenvectors of A. Verify that the sum of the eigenvalues of A is equal to the value of $\mathrm{Tr}(A)$ calculated in (a) and that the three eigenvectors form a basis.

(c) Verify that $U^\dagger A U$ is diagonal and that $U^{-1} = U^\dagger$, where U is the matrix formed by the normalized eigenvectors of A.

(d) Calculate the inverse of $A' = U^\dagger A U$ and verify that A'^{-1} is a diagonal matrix whose eigenvalues are the inverse of those of A'.

Solution

(a) Taking the Hermitian adjoints of the matrices A and B (see (2.221))

$$A^\dagger = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}, \qquad B^\dagger = \begin{pmatrix} 1 & 0 & -i \\ 0 & -2i & 0 \\ 3 & 0 & 5i \end{pmatrix}, \tag{2.445}$$

we see that A is Hermitian and B is not. Using the products

$$AB = \begin{pmatrix} 7 & 0 & 21 \\ 1 & 2i & -5 \\ -i & -2 & 5i \end{pmatrix}, \quad BA = \begin{pmatrix} 7 & 3i & -3 \\ 0 & 2i & 2 \\ 7i & 5 & 5i \end{pmatrix}, \tag{2.446}$$

we can obtain the commutator

$$[A,B]=\begin{pmatrix} 0 & -3i & 24 \\ 1 & 0 & -7 \\ -8i & -7 & 0 \end{pmatrix}. \tag{2.447}$$

From (2.446) we see that

$$\mathrm{Tr}(AB)=7+2i+5i=7+7i=\mathrm{Tr}(BA). \tag{2.448}$$

That is, the cyclic permutation of matrices leaves the trace unchanged; see (2.239). On the other hand, (2.447) shows that the trace of the commutator $[A,\ B]$ is zero: $\mathrm{Tr}([A,\ B])=0+0+0=0$.

(b) The eigenvalues and eigenvectors of A were calculated in Example 2.21 (see (2.302), (2.304), (2.308), (2.310)). We have $a_1=7$, $a_2=\sqrt{2}$, and $a_3=-\sqrt{2}$:

$$|\,a_1\rangle=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\quad |\,a_2\rangle=\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2(2-\sqrt{2})}} \\ \frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} \end{pmatrix},\quad |\,a_3\rangle=\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2(2+\sqrt{2})}} \\ -\frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix}. \tag{2.449}$$

One can easily verify that the eigenvectors $|\,a_1\rangle$, $|\,a_2\rangle$, and $|\,a_3\rangle$ are mutually orthogonal: $\langle a_i\,|\,a_j\rangle=\delta_{ij}$ where $i,j=1,2,3$. Since the set of $|\,a_1\rangle$, $|\,a_2\rangle$, and $|\,a_3\rangle$ satisfy the completeness condition

$$\sum_{j=1}^{3}|\,a_j\rangle\langle a_j\,|=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.450}$$

and since they are orthonormal, they form a complete and orthonormal basis.

(c) The columns of the matrix U are given by the eigenvectors (2.449):

$$U=\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2(2-\sqrt{2})}} & \frac{1}{\sqrt{2(2+\sqrt{2})}} \\ 0 & \frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} & -\frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix}. \tag{2.451}$$

We can show that the product $U^{\dagger}AU$ is diagonal where the diagonal elements are the eigenvalues of the matrix A; $U^{\dagger}AU$ is given by

$$\begin{aligned} &\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2(2-\sqrt{2})}} & -\frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} \\ 0 & \frac{1}{\sqrt{2(2+\sqrt{2})}} & \frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix}\begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2(2-\sqrt{2})}} & \frac{1}{\sqrt{2(2+\sqrt{2})}} \\ 0 & \frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} & -\frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix} \\ &=\begin{pmatrix} 7 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix}. \end{aligned} \tag{2.452}$$

We can also show that $U^\dagger U = 1$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2(2-\sqrt{2})}} & -\frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} \\ 0 & \frac{1}{\sqrt{2(2+\sqrt{2})}} & \frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2(2-\sqrt{2})}} & \frac{1}{\sqrt{2(2+\sqrt{2})}} \\ 0 & \frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} & -\frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.453}$$

This implies that the matrix U is unitary: $U^\dagger = U^{-1}$. Note that, from (2.451), we have $|\det(U)| = |-i| = 1$.

(d) Using (2.452) we can verify that the inverse of $A' = U^\dagger A U$ is a diagonal matrix whose elements are given by the inverse of the diagonal elements of A':

$$A' = \begin{pmatrix} 7 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix} \implies A'^{-1} = \begin{pmatrix} \frac{1}{7} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}. \tag{2.454}$$

Problem 2.11

Consider a particle whose Hamiltonian matrix is $H = \begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

(a) Is $|\lambda\rangle = \begin{pmatrix} i \\ 7i \\ -2 \end{pmatrix}$ an eigenstate of H? Is H Hermitian?

(b) Find the energy eigenvalues, a_1, a_2, and a_3, and the normalized energy eigenvectors, $|a_1\rangle$, $|a_2\rangle$, and $|a_3\rangle$, of H.

(c) Find the matrix corresponding to the operator obtained from the ket-bra product of the first eigenvector $P = |a_1\rangle\langle a_1|$. Is P a projection operator? Calculate the commutator $[P, H]$ firstly by using commutator algebra and then by using matrix products.

Solution

(a) The ket $|\lambda\rangle$ is an eigenstate of H only if the action of the Hamiltonian on $|\lambda\rangle$ is of the form $H|\lambda\rangle = b|\lambda\rangle$, where b is constant. This is not the case here:

$$H|\lambda\rangle = \begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} i \\ 7i \\ -2 \end{pmatrix} = \begin{pmatrix} -7+2i \\ -1+7i \\ 7i \end{pmatrix}. \tag{2.455}$$

Using the definition of the Hermitian adjoint of matrices (2.221), it is easy to ascertain that H is Hermitian:

$$H^\dagger = \begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} = H. \tag{2.456}$$

(b) The energy eigenvalues can be obtained by solving the secular equation

$$\begin{aligned} 0 &= \begin{vmatrix} 2-a & i & 0 \\ -i & 1-a & 1 \\ 0 & 1 & -a \end{vmatrix} = (2-a)[(1-a)(-a)-1] - i(-i)(-a) \\ &= -(a-1)(a-1-\sqrt{3})(a-1+\sqrt{3}), \end{aligned} \tag{2.457}$$

which leads to

$$a_1 = 1, \quad a_2 = 1 - \sqrt{3}, \quad a_3 = 1 + \sqrt{3}. \tag{2.458}$$

To find the eigenvector corresponding to the first eigenvalue, $a_1 = 1$, we need to solve the matrix equation

$$\begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Longrightarrow \begin{array}{rcl} x + iy & = & 0 \\ -ix + z & = & 0 \\ y - z & = & 0 \end{array} \tag{2.459}$$

which yields $x = 1$, $y = z = i$. So the eigenvector corresponding to $a_1 = 1$ is

$$|\, a_1\rangle = \begin{pmatrix} 1 \\ i \\ i \end{pmatrix}. \tag{2.460}$$

This eigenvector is not normalized since $\langle a_1 \,|\, a_1\rangle = 1 + (i^*)(i) + (i^*)(i) = 3$. The normalized $|\, a_1\rangle$ is therefore

$$|\, a_1\rangle = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ i \\ i \end{pmatrix}. \tag{2.461}$$

Solving (2.459) for the other two energy eigenvalues, $a_2 = 1 - \sqrt{3}, a_3 = 1 + \sqrt{3}$, and normalizing, we end up with

$$|\, a_2\rangle = \frac{1}{\sqrt{6(2 - \sqrt{3})}}\begin{pmatrix} i(2 - \sqrt{3}) \\ 1 - \sqrt{3} \\ 1 \end{pmatrix}, \qquad |\, a_3\rangle = \frac{1}{\sqrt{6(2 + \sqrt{3})}}\begin{pmatrix} i(2 + \sqrt{3}) \\ 1 + \sqrt{3} \\ 1 \end{pmatrix}. \tag{2.462}$$

(c) The operator P is given by

$$P = |\, a_1\rangle\langle a_1 \,| = \frac{1}{3}\begin{pmatrix} 1 \\ i \\ i \end{pmatrix}\begin{pmatrix} 1 & -i & -i \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix}. \tag{2.463}$$

Since this matrix is Hermitian and since the square of P is equal to P,

$$P^2 = \frac{1}{9}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix} = P, \tag{2.464}$$

so P is a projection operator. Using the relations $H\,|\, a_1\rangle = |\, a_1\rangle$ and $\langle a_1 \,|\, H = \langle a_1 \,|$ (because H is Hermitian), and since $P = |\, a_1\rangle\langle a_1 \,|$, we can evaluate algebraically the commutator $[P, H]$ as follows:

$$[P, H] = PH - HP = |\, a_1\rangle\langle a_1 \,|\, H - H \,|\, a_1\rangle\langle a_1 \,| = |\, a_1\rangle\langle a_1 \,| - |\, a_1\rangle\langle a_1 \,| = 0. \tag{2.465}$$

We can reach the same result by using the matrices of H and P:

$$\begin{aligned} [P, H] &= \frac{1}{3}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{2.466}$$

Problem 2.12

Consider the matrices $A = \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix}$.

(a) Check if A and B are Hermitian and find the eigenvalues and eigenvectors of A. Any degeneracies?

(b) Verify that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, $\det(AB) = \det(A)\det(B)$, and $\det(B^\dagger) = (\det(B))^*$.

(c) Calculate the commutator $[A, B]$ and the anticommutator $\{A, B\}$.

(d) Calculate the inverses A^{-1}, B^{-1}, and $(AB)^{-1}$. Verify that $(AB)^{-1} = B^{-1}A^{-1}$.

(e) Calculate A^2 and infer the expressions of A^{2n} and A^{2n+1}. Use these results to calculate the matrix of $e^{\alpha A}$.

Solution

(a) The matrix A is Hermitian but B is not. The eigenvalues of A are $a_1 = -1$ and $a_2 = a_3 = 1$ and its normalized eigenvectors are

$$|a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix}, \qquad |a_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -i \end{pmatrix}, \qquad |a_3\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{2.467}$$

Note that the eigenvalue 1 is doubly degenerate, since the two eigenvectors $|a_2\rangle$ and $|a_3\rangle$ correspond to the same eigenvalue $a_2 = a_3 = 1$.

(b) A calculation of the products (AB) and (BA) reveals that the traces $\mathrm{Tr}(AB)$ and $\mathrm{Tr}(BA)$ are equal:

$$\begin{aligned} \mathrm{Tr}(AB) &= \mathrm{Tr}\begin{pmatrix} 0 & 1 & -2i \\ 3 & 1 & 5 \\ -2i & 1 & 0 \end{pmatrix} = 1, \\ \mathrm{Tr}(BA) &= \mathrm{Tr}\begin{pmatrix} 0 & i & 2i \\ -5i & 1 & 3i \\ 2i & -i & 0 \end{pmatrix} = 1 = \mathrm{Tr}(AB). \end{aligned} \tag{2.468}$$

From the matrices A and B, we have $\det(A) = i(i) = -1$, $\det(B) = -4 + 16i$. We can thus write

$$\det(AB) = \det\begin{pmatrix} 0 & 1 & -2i \\ 3 & 1 & 5 \\ -2i & 1 & 0 \end{pmatrix} = 4 - 16i = (-1)(-4 + 16i) = \det(A)\det(B). \tag{2.469}$$

On the other hand, since $\det(B) = -4 + 16i$ and $\det(B^\dagger) = -4 - 16i$, we see that $\det(B^\dagger) = -4 - 16i = (-4 + 16i)^* = (\det(B))^*$.

(c) The commutator $[A, B]$ is given by

$$AB - BA = \begin{pmatrix} 0 & 1 & -2i \\ 3 & 1 & 5 \\ -2i & 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & i & 2i \\ -5i & 1 & 3i \\ 2i & -i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1-i & -4i \\ 3+5i & 0 & 5-3i \\ -4i & 1+i & 0 \end{pmatrix} \tag{2.470}$$

and the anticommutator $\{A, B\}$ by

$$AB + BA = \begin{pmatrix} 0 & 1 & -2i \\ 3 & 1 & 5 \\ -2i & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & i & 2i \\ -5i & 1 & 3i \\ 2i & -i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1+i & 0 \\ 3-5i & 2 & 5+3i \\ 0 & 1-i & 0 \end{pmatrix}. \tag{2.471}$$

(d) A calculation similar to (2.233) leads to the inverses of A, B, and AB:

$$A^{-1}=\begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}, \qquad B^{-1}=\frac{1}{68}\begin{pmatrix} 22+3i & 8-2i & 20-5i \\ -6-24i & 4+16i & 10+40i \\ -12+3i & 8-2i & -14-5i \end{pmatrix}, \tag{2.472}$$

$$(AB)^{-1}=\frac{1}{68}\begin{pmatrix} -5-20i & 8-2i & -3+22i \\ 40-10i & 4+16i & 24-6i \\ -5+14i & 8-2i & -3-12i \end{pmatrix}. \tag{2.473}$$

From (2.472) it is now easy to verify that the product $B^{-1}A^{-1}$ is equal to $(AB)^{-1}$:

$$B^{-1}A^{-1}=\frac{1}{68}\begin{pmatrix} -5-20i & 8-2i & -3+22i \\ 40-10i & 4+16i & 24-6i \\ -5+14i & 8-2i & -3-12i \end{pmatrix}=(AB)^{-1}. \tag{2.474}$$

(e) Since

$$A^2=\begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}=I, \tag{2.475}$$

we can write $A^3=A$, $A^4=I$, $A^5=A$, and so on. We can generalize these relations to any value of n: $A^{2n}=I$ and $A^{2n+1}=A$:

$$A^{2n}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}=I, \qquad A^{2n+1}=\begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}=A. \tag{2.476}$$

Since $A^{2n}=I$ and $A^{2n+1}=A$, we can write

$$e^{xA}=\sum_{n=0}^{\infty}\frac{x^nA^n}{n!}=\sum_{n=0}^{\infty}\frac{x^{2n}A^{2n}}{(2n)!}+\sum_{n=0}^{\infty}\frac{x^{2n+1}A^{2n+1}}{(2n+1)!}=I\sum_{n=0}^{\infty}\frac{x^{2n}}{(2n)!}+A\sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!}. \tag{2.477}$$

The relations

$$\sum_{n=0}^{\infty}\frac{x^{2n}}{(2n)!}=\cosh x, \qquad \sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!}=\sinh x, \tag{2.478}$$

lead to

$$\begin{aligned} e^{xA} &= I\cosh x+A\sinh x=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\cosh x+\begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}\sinh x \\ &= \begin{pmatrix} \cosh x & 0 & i\sinh x \\ 0 & \cosh x+\sinh x & 0 \\ -i\sinh x & 0 & \cosh x \end{pmatrix}. \end{aligned} \tag{2.479}$$

Problem 2.13

Consider two matrices: $A=\begin{pmatrix} 0 & i & 2 \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}$ and $B=\begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix}$. Calculate $A^{-1}B$ and BA^{-1}. Are they equal?

Solution

As mentioned above, a calculation similar to (2.233) leads to the inverse of A:

$$A^{-1} = \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ 1/2 & -i/2 & 0 \end{pmatrix}. \tag{2.480}$$

The products $A^{-1}B$ and BA^{-1} are given by

$$A^{-1}B = \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ 1/2 & -i/2 & 0 \end{pmatrix}\begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -2i \\ 3 & 1 & 5 \\ 1-3i/2 & 0 & -5i/2 \end{pmatrix}, \tag{2.481}$$

$$BA^{-1} = \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix}\begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ 1/2 & -i/2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & i & 2i \\ 5/2 & 1-5i/2 & 3i \\ -1 & 0 & 0 \end{pmatrix}. \tag{2.482}$$

We see that $A^{-1}B$ and BA^{-1} are not equal.

Remark

We should note that the quotient B/A of two matrices A and B is equal to the product BA^{-1} and not $A^{-1}B$; that is:

$$\frac{B}{A} = BA^{-1} = \frac{\begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix}}{\begin{pmatrix} 0 & i & 2 \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}} = \begin{pmatrix} 0 & i & 2i \\ 5/2 & 1-5i/2 & 3i \\ -1 & 0 & 0 \end{pmatrix}. \tag{2.483}$$

Problem 2.14

Consider the matrices $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

(a) Find the eigenvalues and normalized eigenvectors of A and B. Denote the eigenvectors of A by $|a_1\rangle, |a_2\rangle, |a_3\rangle$ and those of B by $|b_1\rangle, |b_2\rangle, |b_3\rangle$. Are there any degenerate eigenvalues?

(b) Show that each of the sets $|a_1\rangle, |a_2\rangle, |a_3\rangle$ and $|b_1\rangle, |b_2\rangle, |b_3\rangle$ forms an orthonormal and complete basis, i.e., show that $\langle a_j | a_k\rangle = \delta_{jk}$ and $\sum_{j=1}^{3} |a_j\rangle\langle a_j| = I$, where I is the 3×3 unit matrix; then show that the same holds for $|b_1\rangle, |b_2\rangle, |b_3\rangle$.

(c) Find the matrix U of the transformation from the basis $\{|a\rangle\}$ to $\{|b\rangle\}$. Show that $U^{-1} = U^{\dagger}$. Verify that $U^{\dagger}U = I$. Calculate how the matrix A transforms under U, i.e., calculate $A' = UAU^{\dagger}$.

Solution

(a) It is easy to verify that the eigenvalues of A are $a_1 = 0$, $a_2 = \sqrt{2}$, $a_3 = -\sqrt{2}$ and their corresponding normalized eigenvectors are

$$|a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \qquad |a_2\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}, \qquad |a_3\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}. \tag{2.484}$$

The eigenvalues of B are $b_1 = 1$, $b_2 = 0$, $b_3 = -1$ and their corresponding normalized eigenvectors are

$$|\, b_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |\, b_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |\, b_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.485}$$

None of the eigenvalues of A and B are degenerate.

(b) The set $|\, a_1\rangle, |\, a_2\rangle, |\, a_3\rangle$ is indeed complete because the sum of $|\, a_1\rangle\langle a_1\,|$, $|\, a_2\rangle\langle a_2\,|$, and $|\, a_3\rangle\langle a_3\,|$ as given by

$$|\, a_1\rangle\langle a_1\,| = \frac{1}{2}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \tag{2.486}$$

$$|\, a_2\rangle\langle a_2\,| = \frac{1}{4}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}\begin{pmatrix} 1 & \sqrt{2} & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 2 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}, \tag{2.487}$$

$$|\, a_3\rangle\langle a_3\,| = \frac{1}{4}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}\begin{pmatrix} 1 & -\sqrt{2} & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & -\sqrt{2} & 1 \\ -\sqrt{2} & 2 & -\sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{pmatrix}, \tag{2.488}$$

is equal to unity:

$$\begin{aligned} \sum_{j=1}^{3} |\, a_j\rangle\langle a_j\,| &= \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 2 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix} \\ &\quad + \frac{1}{4}\begin{pmatrix} 1 & -\sqrt{2} & 1 \\ -\sqrt{2} & 2 & -\sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{2.489}$$

The states $|\, a_1\rangle, |\, a_2\rangle, |\, a_3\rangle$ are orthonormal, since $\langle a_1 \mid a_2\rangle = \langle a_1 \mid a_3\rangle = \langle a_3 \mid a_2\rangle = 0$ and $\langle a_1 \mid a_1\rangle = \langle a_2 \mid a_2\rangle = \langle a_3 \mid a_3\rangle = 1$. Following the same procedure, we can ascertain that

$$|\, b_1\rangle\langle b_1\,| + |\, b_2\rangle\langle b_2\,| + |\, b_3\rangle\langle b_3\,| = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.490}$$

We can verify that the states $|\, b_1\rangle, |\, b_2\rangle, |\, b_3\rangle$ are orthonormal, since $\langle b_1 \mid b_2\rangle = \langle b_1 \mid b_3\rangle = \langle b_3 \mid b_2\rangle = 0$ and $\langle b_1 \mid b_1\rangle = \langle b_2 \mid b_2\rangle = \langle b_3 \mid b_3\rangle = 1$.

(c) The elements of the matrix U, corresponding to the transformation from the basis $\{|\, a\rangle\}$ to $\{|\, b\rangle\}$, are given by $U_{jk} = \langle b_j \mid a_k\rangle$ where $j, k = 1,2,3$:

$$U = \begin{pmatrix} \langle b_1 \mid a_1\rangle & \langle b_1 \mid a_2\rangle & \langle b_1 \mid a_3\rangle \\ \langle b_2 \mid a_1\rangle & \langle b_2 \mid a_2\rangle & \langle b_2 \mid a_3\rangle \\ \langle b_3 \mid a_1\rangle & \langle b_3 \mid a_2\rangle & \langle b_3 \mid a_3\rangle \end{pmatrix}, \tag{2.491}$$

where the elements $\langle b_j \mid a_k \rangle$ can be calculated from (2.484) and (2.485):

$$U_{11} = \langle b_1 \mid a_1 \rangle = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = -\frac{\sqrt{2}}{2}, \tag{2.492}$$

$$U_{12} = \langle b_1 \mid a_2 \rangle = \tfrac{1}{2} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} = \frac{1}{2}, \tag{2.493}$$

$$U_{13} = \langle b_1 \mid a_3 \rangle = \tfrac{1}{2} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} = \frac{1}{2}, \tag{2.494}$$

$$U_{21} = \langle b_2 \mid a_1 \rangle = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = 0, \tag{2.495}$$

$$U_{22} = \langle b_2 \mid a_2 \rangle = \tfrac{1}{2} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} = \frac{\sqrt{2}}{2}, \tag{2.496}$$

$$U_{23} = \langle b_2 \mid a_3 \rangle = \tfrac{1}{2} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} = -\frac{\sqrt{2}}{2}, \tag{2.497}$$

$$U_{31} = \langle b_3 \mid a_1 \rangle = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \frac{\sqrt{2}}{2}, \tag{2.498}$$

$$U_{32} = \langle b_3 \mid a_2 \rangle = \tfrac{1}{2} \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} = \frac{1}{2}, \tag{2.499}$$

$$U_{33} = \langle b_3 \mid a_3 \rangle = \tfrac{1}{2} \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} = \frac{1}{2}. \tag{2.500}$$

Collecting these elements, we obtain

$$U = \frac{1}{2} \begin{pmatrix} -\sqrt{2} & 1 & 1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1 & 1 \end{pmatrix}. \tag{2.501}$$

Calculating the inverse of U as we did in (2.233), we see that it is equal to its Hermitian adjoint:

$$U^{-1} = \frac{1}{2} \begin{pmatrix} -\sqrt{2} & 0 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \\ 1 & -\sqrt{2} & 1 \end{pmatrix} = U^{\dagger}. \tag{2.502}$$

This implies that the matrix U is unitary. The matrix A transforms as follows:

$$A' = UAU^\dagger = \frac{1}{4}\begin{pmatrix} -\sqrt{2} & 1 & 1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} -\sqrt{2} & 0 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \\ 1 & -\sqrt{2} & 1 \end{pmatrix}$$
$$= \frac{1}{2}\begin{pmatrix} 1-\sqrt{2} & -1 & 1 \\ -1 & -2 & 1 \\ 1 & 1 & 1+\sqrt{2} \end{pmatrix}. \tag{2.503}$$

Problem 2.15
Show that the product of an antiunitary operator, $\hat{F}$, times a unitary operator, $\hat{U}$, is an antiunitary operator.

Solution
Let us denote the product of $\hat{F}$ and $\hat{U}$ by $\hat{T} = \hat{F}\hat{U}$ and the transforms of any two states $|\psi_1\rangle$ and $|\psi_2\rangle$ by $|\psi_1'\rangle$ and $|\psi_2'\rangle$, respectively:

$$\hat{T}|\psi_1\rangle = |\psi_1'\rangle, \qquad \hat{T}|\psi_2\rangle = |\psi_2'\rangle. \tag{2.504}$$

Hence, for any states $|\psi_1\rangle$ and $|\psi_2\rangle$ and any complex numbers a_1 and a_2, we can write

$$\begin{aligned} \hat{T}(a_1|\psi_1\rangle + a_2|\psi_2\rangle) &= \hat{F}\hat{U}(a_1|\psi_1\rangle) + \hat{F}\hat{U}(a_2|\psi_2\rangle) = \hat{F}(a_1\hat{U}|\psi_1\rangle) + \hat{F}(a_2\hat{U}|\psi_2\rangle) \\ &= a_1^*\hat{F}\hat{U}|\psi_1\rangle + a_2^*\hat{F}\hat{U}|\psi_2 \\ &= a_1^*\hat{T}|\psi_1\rangle + a_2^*\hat{T}|\psi_2 \\ &= a_1^*|\psi_1'\rangle + a_2^*|\psi_2'\rangle, \end{aligned} \tag{2.505}$$

where we used the fact that unitary operators commute with complex numbers in the first line; i.e., $\hat{U}a_j = a_j\hat{U}$ where $j = 1, 2$. Equations (2.505) shows the operator $\hat{T}$ is antilinear. Additionally, the scalar product $\langle\psi_1|\psi_2\rangle$ transforms under $\hat{T}$ as follows:

$$\begin{aligned} \langle\psi_1'|\psi_2'\rangle &= \left(\langle\psi_1|\hat{T}^\dagger\right)\left(\hat{T}|\psi_2\rangle\right) = \\ &= \left(\langle\psi_1|\hat{U}^\dagger\right)\hat{F}^\dagger\hat{F}\left(\hat{U}|\psi_2\rangle\right) \\ &= \left(\langle\psi_1|\hat{U}^\dagger\hat{U}|\psi_2\rangle\right)^* \\ &= \langle\psi_1|\psi_2\rangle^* \\ &= \langle\psi_2|\psi_1\rangle, \end{aligned} \tag{2.506}$$

where we have used the relation $\hat{U}^\dagger\hat{U} = I$ in the fourth line.

From (2.505) and (2.506), we conclude that the operator $\hat{T} = \hat{F}\hat{U}$ is both antilinear and antiunitary. That is, *the product of an antiunitary operator by a unitary operator is an antiunitary operator.*

Problem 2.16
Calculate the following expressions involving Dirac's delta function:

(a) $\int_{-5}^{5} \cos(3x)\delta(x-\pi/3)\,dx$

(b) $\int_{0}^{10} \left[e^{2x-7}+4\right]\delta(x+3)\,dx$

(c) $\left[2\cos^2(3x)-\sin(x/2)\right]\delta(x+\pi)$

(d) $\int_{0}^{\pi} \cos(3\theta)\delta'''(\theta-\pi/2)\,d\theta$

(e) $\int_{2}^{9} \left(x^2-5x+2\right)\delta[2(x-4)]\,dx$.

Solution

(a) Since $x=\pi/3$ lies within the interval $(-5,\ 5)$, equation (2.317) yields

$$\int_{-5}^{5} \cos(3x)\delta(x-\pi/3)\,dx = \cos\left(3\frac{\pi}{3}\right) = -1. \tag{2.507}$$

(b) Since $x=-3$ lies outside the interval $(0,\ 10)$, Eq (2.317) yields at once

$$\int_{0}^{10} \left[e^{2x-7}+4\right]\delta(x+3)\,dx = 0. \tag{2.508}$$

(c) Using the relation $f(x)\delta(x-a)=f(a)\delta(x-a)$ which is listed in Appendix A, we have

$$\begin{aligned}\left[2\cos^2(3x)-\sin(x/2)\right]\delta(x+\pi) &= \left[2\cos^2(3(-\pi))-\sin((-\pi)/2)\right]\delta(x+\pi)\\ &= 3\delta(x+\pi).\end{aligned} \tag{2.509}$$

(d) Inserting $n=3$ into Eq (2.318) and since $\cos'''(3\theta)=27\sin(3\theta)$, we obtain

$$\begin{aligned}\int_{0}^{\pi} \cos(3\theta)\delta'''(\theta-\pi/2)\,d\theta &= (-1)^3\cos'''(3\pi/2) = (-1)^3 27\sin(3\pi/2)\\ &= 27.\end{aligned} \tag{2.510}$$

(e) Since $\delta[2(x-4)]=(1/2)\delta(x-4)$, we have

$$\begin{aligned}\int_{2}^{9} \left(x^2-5x+2\right)\delta[2(x-4)],dx &= \frac{1}{2}\int_{2}^{9} \left(x^2-5x+2\right)\delta(x-4)\,dx\\ &= \frac{1}{2}\left(4^2-5\times 4+2\right) = -1.\end{aligned} \tag{2.511}$$

Problem 2.17

Consider a system whose Hamiltonian is given by $\hat{H}=\alpha\left(|\,\phi_1\rangle\langle\phi_2\,|+|\,\phi_2\rangle\langle\phi_1\,|\right)$, where α is a real number having the dimensions of energy and $|\,\phi_1\rangle$, $|\,\phi_2\rangle$ are normalized eigenstates of a Hermitian operator $\hat{A}$ that has no degenerate eigenvalues.

(a) Is $\hat{H}$ a projection operator? What about $\alpha^{-2}\hat{H}^2$?

(b) Show that $|\,\phi_1\rangle$ and $|\,\phi_2\rangle$ are not eigenstates of $\hat{H}$.

(c) Calculate the commutators $[\hat{H},\ |\,\phi_1\rangle\langle\phi_1\,|]$ and $[\hat{H},\ |\,\phi_2\rangle\langle\phi_2\,|]$ then find the relation that may exist between them.

(d) Find the normalized eigenstates of $\hat{H}$ and their corresponding energy eigenvalues.

(e) Assuming that $|\,\phi_1\rangle$ and $|\,\phi_2\rangle$ form a complete and orthonormal basis, find the matrix representing $\hat{H}$ in the basis. Find the eigenvalues and eigenvectors of the matrix and compare the results with those derived in (d).

Solution

(a) Since $|\phi_1\rangle$ and $|\phi_2\rangle$ are eigenstates of $\hat{A}$ and since $\hat{A}$ is Hermitian, they must be orthogonal, $\langle\phi_1|\phi_2\rangle = 0$ (instance of Theorem 2.1). Now, since $|\phi_1\rangle$ and $|\phi_2\rangle$ are both normalized and since $\langle\phi_1|\phi_2\rangle = 0$, we can reduce $\hat{H}^2$ to

$$\begin{aligned}\hat{H}^2 &= \alpha^2\big(|\phi_1\rangle\langle\phi_2| + |\phi_2\rangle\langle\phi_1|\big)\big(|\phi_1\rangle\langle\phi_2| + |\phi_2\rangle\langle\phi_1|\big)\\ &= \alpha^2\big(|\phi_1\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2|\big),\end{aligned} \tag{2.512}$$

which is different from $\hat{H}$; hence $\hat{H}$ is not a projection operator. The operator $\alpha^{-2}\hat{H}^2$ is a projection operator since it is both Hermitian and equal to its own square. Using (2.512) we can write

$$\begin{aligned}(\alpha^{-2}\hat{H}^2)^2 &= \big(|\phi_1\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2|\big)\big(|\phi_1\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2|\big)\\ &= |\phi_1\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2| = \alpha^{-2}\hat{H}^2.\end{aligned} \tag{2.513}$$

(b) Since $|\phi_1\rangle$ and $|\phi_2\rangle$ are both normalized, and since $\langle\phi_1|\phi_2\rangle = 0$, we have

$$\hat{H}|\phi_1\rangle = \alpha|\phi_1\rangle\langle\phi_2|\phi_1\rangle + \alpha|\phi_2\rangle\langle\phi_1|\phi_1\rangle = \alpha|\phi_2\rangle, \tag{2.514}$$

$$\hat{H}|\phi_2\rangle = \alpha|\phi_1\rangle; \tag{2.515}$$

hence $|\phi_1\rangle$ and $|\phi_2\rangle$ are not eigenstates of $\hat{H}$. In addition, we have

$$\langle\phi_1|\hat{H}|\phi_1\rangle = \langle\phi_2|\hat{H}|\phi_2\rangle = 0. \tag{2.516}$$

(c) Using the relations derived above, $\hat{H}|\phi_1\rangle = \alpha|\phi_2\rangle$ and $\hat{H}|\phi_2\rangle = \alpha|\phi_1\rangle$, we can write

$$[\hat{H}, |\phi_1\rangle\langle\phi_1|] = \alpha\big(|\phi_2\rangle\langle\phi_1| - |\phi_1\rangle\langle\phi_2|\big), \tag{2.517}$$

$$[\hat{H}, |\phi_2\rangle\langle\phi_2|] = \alpha\big(|\phi_1\rangle\langle\phi_2| - |\phi_2\rangle\langle\phi_1|\big); \tag{2.518}$$

hence

$$[\hat{H}, |\phi_1\rangle\langle\phi_1|] = -[\hat{H}, |\phi_2\rangle\langle\phi_2|]. \tag{2.519}$$

(d) Consider a general state $|\psi\rangle = \lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle$. Applying $\hat{H}$ to this state, we get

$$\begin{aligned}\hat{H}|\psi\rangle &= \alpha\big(|\phi_1\rangle\langle\phi_2| + |\phi_2\rangle\langle\phi_1|\big)\big(\lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle\big)\\ &= \alpha\big(\lambda_2|\phi_1\rangle + \lambda_1|\phi_2\rangle\big).\end{aligned} \tag{2.520}$$

Now, since $|\psi\rangle$ is normalized, we have

$$\langle\psi|\psi\rangle = |\lambda_1|^2 + |\lambda_2|^2 = 1. \tag{2.521}$$

The previous two equations show that $|\lambda_1| = |\lambda_2| = 1/\sqrt{2}$ and that $\lambda_1 = \pm\lambda_2$. Hence the eigenstates of the system are:

$$|\psi_\pm\rangle = \frac{1}{\sqrt{2}}\big(|\phi_1\rangle \pm |\phi_2\rangle\big). \tag{2.522}$$

The corresponding eigenvalues are $\pm\alpha$:

$$\hat{H} \mid \psi_{\pm}\rangle = \pm\alpha \mid \psi_{\pm}\rangle. \tag{2.523}$$

(e) Since $\langle\phi_1 \mid \phi_2\rangle = \langle\phi_2 \mid \phi_1\rangle = 0$ and $\langle\phi_1 \mid \phi_1\rangle = \langle\phi_2 \mid \phi_2\rangle = 1$, we can verify that $H_{11} = \langle\phi_1 \mid \hat{H} \mid \phi_1\rangle = 0, H_{22} = \langle\phi_2 \mid \hat{H} \mid \phi_2\rangle = 0, H_{12} = \langle\phi_1 \mid \hat{H} \mid \phi_2\rangle = \alpha, H_{21} = \langle\phi_2 \mid \hat{H} \mid \phi_1\rangle = \alpha$. The matrix of $\hat{H}$ is thus given by

$$H = \alpha \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.524}$$

The eigenvalues of this matrix are equal to $\pm\alpha$ and the corresponding eigenvectors are $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$. These results are indeed similar to those derived in (d).

Problem 2.18

Consider the matrices $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & -3i \\ 0 & 3i & 5 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & -i & 3i \\ -i & 0 & i \\ 3i & i & 0 \end{pmatrix}$.

(a) Check the hermiticity of A and B.

(b) Find the eigenvalues of A and B; denote the eigenvalues of A by a_1, a_2, and a_3. Explain why the eigenvalues of A are real and those of B are imaginary.

(c) Calculate $\mathrm{Tr}(A)$ and $\det(A)$. Verify $\mathrm{Tr}(A) = a_1 + a_2 + a_3$, $\det(A) = a_1 a_2 a_3$.

Solution

(a) Matrix A is Hermitian but B is anti-Hermitian:

$$A^{\dagger} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & -3i \\ 0 & 3i & 5 \end{pmatrix} = A, \qquad B^{\dagger} = \begin{pmatrix} 0 & i & -3i \\ i & 0 & -i \\ -3i & -i & 0 \end{pmatrix} = -B. \tag{2.525}$$

(b) The eigenvalues of A are $a_1 = 6 - \sqrt{10}$, $a_2 = 1$, and $a_3 = 6 + \sqrt{10}$ and those of B are $b_1 = -i\left(3 + \sqrt{17}\right)/2$, $b_2 = 3i$, and $b_3 = i\left(-3 + \sqrt{17}\right)/2$. The eigenvalues of A are real and those of B are imaginary. This is expected since, as shown in (2.87) and (2.88), the expectation values of Hermitian operators are real and those of anti-Hermitian operators are imaginary.

(c) A direct calculation of the trace and the determinant of A yields $\mathrm{Tr}(A) = 1 + 7 + 5 = 13$ and $\det(A) = (7)(5) - (3i)(-3i) = 26$. Adding and multiplying the eigenvalues $a_1 = 6 - \sqrt{10}$, $a_2 = 1$, $a_3 = 6 + \sqrt{10}$, we have $a_1 + a_2 + a_3 = 6 - \sqrt{10} + 1 + 6 + \sqrt{10} = 13$ and $a_1 a_2 a_3 = (6 - \sqrt{10})(1)(6 + \sqrt{10}) = 26$. This confirms the results (2.294) and (2.295):

$$\mathrm{Tr}(A) = a_1 + a_2 + a_3 = 13, \qquad \det(A) = a_1 a_2 a_3 = 26. \tag{2.526}$$

Problem 2.19

Consider a one-dimensional particle which moves along the x-axis and whose Hamiltonian is $\hat{H} = -\mathcal{E}d^2/dx^2 + 16\mathcal{E}\hat{X}^2$, where $\mathcal{E}$ is a real constant having the dimensions of energy.

(a) Is $\psi(x) = Ae^{-2x^2}$, where A is a normalization constant that needs to be found, an eigenfunction of $\hat{H}$? If yes, find the energy eigenvalue.

(b) Calculate the probability of finding the particle anywhere along the negative x-axis.

(c) Find the energy eigenvalue corresponding to the wave function $\phi(x) = 2x\psi(x)$.

(d) Specify the parities of $\phi(x)$ and $\psi(x)$. Are $\phi(x)$ and $\psi(x)$ orthogonal?

Solution

(a) The integral $\int_{-\infty}^{+\infty} e^{-4x^2}dx = \sqrt{\pi}/2$ allows us to find the normalization constant:

$$1 = \int_{-\infty}^{+\infty} |\psi(x)|^2\,dx = A^2 \int_{-\infty}^{+\infty} e^{-4x^2}dx = A^2\frac{\sqrt{\pi}}{2}; \tag{2.527}$$

this leads to $A = \sqrt{2/\sqrt{\pi}}$ and hence $\psi(x) = \sqrt{2/\sqrt{\pi}}e^{-2x^2}$. Since the first and second derivatives of $\psi(x)$ are given by

$$\psi'(x) = \frac{d\psi(x)}{dx} = -4x\psi(x), \qquad \psi''(x) = \frac{d^2\psi(x)}{dx^2} = (16x^2 - 4)\psi(x), \tag{2.528}$$

we see that $\psi(x)$ is an eigenfunction of $\hat{H}$ with an energy eigenvalue equal to $4\mathcal{E}$:

$$\hat{H}\psi(x) = -\mathcal{E}\frac{d^2\psi(x)}{dx^2} + 16\mathcal{E}x^2\psi(x) = -\mathcal{E}(16x^2 - 4)\psi(x) + 16\mathcal{E}x^2\psi(x) = 4\mathcal{E}\psi(x). \tag{2.529}$$

(b) Since $\int_{-\infty}^{0} e^{-4x^2}dx = \sqrt{\pi}/4$, the probability of finding the particle anywhere along the negative x-axis is equal to $\frac{1}{2}$:

$$\int_{-\infty}^{0} |\psi(x)|^2\,dx = \frac{2}{\sqrt{\pi}}\int_{-\infty}^{0} e^{-4x^2}dx = \frac{1}{2}. \tag{2.530}$$

This is expected, since this probability is half the total probability, which in turn is equal to one.

(c) Since the second derivative of $\phi(x) = 2x\psi(x)$ is $\phi''(x) = 4\psi'(x) + 2x\psi''(x) = 8x(-3 + 4x^2)\psi(x) = 4(-3+4x^2)\phi(x)$, we see that $\phi(x)$ is an eigenfunction of $\hat{H}$ with an energy eigenvalue equal to $12\mathcal{E}$:

$$\hat{H}\phi(x) = -\mathcal{E}\frac{d^2\phi(x)}{dx^2} + 16\mathcal{E}x^2\phi(x) = -4\mathcal{E}(-3 + 4x^2)\phi(x) + 16\mathcal{E}x^2\phi(x) = 12\mathcal{E}\phi(x). \tag{2.531}$$

(d) The wave functions $\psi(x)$ and $\phi(x)$ are even and odd, respectively, since $\psi(-x) = \psi(x)$ and $\phi(-x) = -\phi(x)$; hence their product is an odd function. Therefore, they are orthogonal, since the symmetric integration of an odd function is zero:

$$\begin{aligned}\langle\phi \mid \psi\rangle &= \int_{-\infty}^{+\infty} \phi^*(x)\psi(x)\,dx = \int_{-\infty}^{+\infty} \phi(x)\psi(x)\,dx = \int_{+\infty}^{-\infty} \phi(-x)\psi(-x)(-dx) \\ &= -\int_{-\infty}^{+\infty} \phi(x)\psi(x)\,dx = 0.\end{aligned} \tag{2.532}$$

Problem 2.20

(a) Find the eigenvalues and the eigenfunctions of the operator $\hat{A} = -d^2/dx^2$; restrict the search for the eigenfunctions to those complex functions that vanish everywhere except in the region $0 < x < a$.

(b) Normalize the eigenfunction and find the probability in the region $0 < x < a/2$.

Solution

(a) The eigenvalue problem for $-d^2/dx^2$ consists of solving the differential equation

$$-\frac{d^2\psi(x)}{dx^2} = \alpha\psi(x) \tag{2.533}$$

and finding the eigenvalues α and the eigenfunction $\psi(x)$. The most general solution to this equation is

$$\psi(x) = Ae^{ibx} + Be^{-ibx}, \tag{2.534}$$

with $\alpha = b^2$. Using the boundary conditions of $\psi(x)$ at $x = 0$ and $x = a$, we have

$$\psi(0) = A + B = 0 \implies B = -A, \qquad \psi(a) = Ae^{iba} + Be^{-iba} = 0. \tag{2.535}$$

A substitution of $B = -A$ into the second equation leads to $A\left(e^{iba} - e^{-iba}\right) = 0$ or $e^{iba} = e^{-iba}$ which leads to $e^{2iba} = 1$. Thus, we have $\sin 2ba = 0$ and $\cos 2ba = 1$, so $ba = n\pi$. The eigenvalues are then given by $\alpha_n = n^2\pi^2/a^2$ and the corresponding eigenvectors by $\psi_n(x) = A\left(e^{in\pi x/a} - e^{-in\pi x/a}\right)$; that is,

$$\alpha_n = \frac{n^2\pi^2}{a^2}, \qquad \psi_n(x) = C_n \sin\left(\frac{n\pi x}{a}\right). \tag{2.536}$$

So the eigenvalue spectrum of the operator $\hat{A} = -d^2/dx^2$ is discrete, because the eigenvalues and eigenfunctions depend on a discrete number n.

(b) The normalization of $\psi_n(x)$,

$$1 = C_n^2 \int_0^a \sin^2\left(\frac{n\pi x}{a}\right) dx = \frac{C_n^2}{2}\int_0^a \left[1 - \cos\left(\frac{2n\pi x}{a}\right)\right] dx = \frac{C_n^2}{2}a, \tag{2.537}$$

yields $C_n = \sqrt{2/a}$ and hence $\psi_n(x) = \sqrt{2/a}\sin(n\pi x/a)$. The probability in the region $0 < x < a/2$ is given by

$$\frac{2}{a}\int_0^{a/2} \sin^2\left(\frac{n\pi x}{a}\right) dx = \frac{1}{a}\int_0^{a/2} \left[1 - \cos\left(\frac{2n\pi x}{a}\right)\right] dx = \frac{1}{2}. \tag{2.538}$$

This is expected since the total probability is 1: $\int_0^a |\psi_n(x)|^2\, dx = 1$.

2.10 Exercises

Exercise 2.1

Consider the two states $|\psi\rangle = i|\phi_1\rangle + 3i|\phi_2\rangle - |\phi_3\rangle$ and $|\chi\rangle = |\phi_1\rangle - i|\phi_2\rangle + 5i|\phi_3\rangle$, where $|\phi_1\rangle$, $|\phi_2\rangle$ and $|\phi_3\rangle$ are orthonormal.

(a) Calculate $\langle\psi|\psi\rangle$, $\langle\chi|\chi\rangle$, $\langle\psi|\chi\rangle$, $\langle\chi|\psi\rangle$, and infer $\langle\psi + \chi|\psi + \chi\rangle$. Are the scalar products $\langle\psi|\chi\rangle$ and $\langle\chi|\psi\rangle$ equal?

(b) Calculate $|\psi\rangle\langle\chi|$ and $|\chi\rangle\langle\psi|$. Are they equal? Calculate their traces and compare them.

(c) Find the Hermitian conjugates of $|\psi\rangle$, $|\chi\rangle$, $|\psi\rangle\langle\chi|$, and $|\chi\rangle\langle\psi|$.

Exercise 2.2
Consider two states $|\psi_1\rangle = |\phi_1\rangle + 4i|\phi_2\rangle + 5|\phi_3\rangle$ and $|\psi_2\rangle = b|\phi_1\rangle + 4|\phi_2\rangle - 3i|\phi_3\rangle$, where $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ are orthonormal kets, and where b is a constant. Find the value of b so that $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthogonal.

Exercise 2.3
If $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ are orthonormal, show that the states $|\psi\rangle = i|\phi_1\rangle + 3i|\phi_2\rangle - |\phi_3\rangle$ and $|\chi\rangle = |\phi_1\rangle - i|\phi_2\rangle + 5i|\phi_3\rangle$ satisfy
(a) the triangle inequality and
(b) the Schwarz inequality.

Exercise 2.4
Find the constant α so that the states $|\psi\rangle = \alpha|\phi_1\rangle + 5|\phi_2\rangle$ and $|\chi\rangle = 3\alpha|\phi_1\rangle - 4|\phi_2\rangle$ are orthogonal; consider $|\phi_1\rangle$ and $|\phi_2\rangle$ to be orthonormal.

Exercise 2.5
If $|\psi\rangle = |\phi_1\rangle + |\phi_2\rangle$ and $|\chi\rangle = |\phi_1\rangle - |\phi_2\rangle$, prove the following relations (note that $|\phi_1\rangle$ and $|\phi_2\rangle$ are not orthonormal):
(a) $\langle\psi|\psi\rangle + \langle\chi|\chi\rangle = 2\langle\phi_1|\phi_1\rangle + 2\langle\phi_2|\phi_2\rangle$,
(b) $\langle\psi|\psi\rangle - \langle\chi|\chi\rangle = 2\langle\phi_1|\phi_2\rangle + 2\langle\phi_2|\phi_1\rangle$.

Exercise 2.6
Consider a state which is given in terms of three orthonormal vectors $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ as follows:

$$|\psi\rangle = \frac{1}{\sqrt{15}}|\phi_1\rangle + \frac{1}{\sqrt{3}}|\phi_2\rangle + \frac{1}{\sqrt{5}}|\phi_3\rangle,$$

where $|\phi_n\rangle$ are eigenstates to an operator $\hat{B}$ such that: $\hat{B}|\phi_n\rangle = (3n^2 - 1)|\phi_n\rangle$ with $n = 1, 2, 3$.
(a) Find the norm of the state $|\psi\rangle$.
(b) Find the expectation value of $\hat{B}$ for the state $|\psi\rangle$.
(c) Find the expectation value of $\hat{B}^2$ for the state $|\psi\rangle$.

Exercise 2.7
Are the following sets of functions linearly independent or dependent?
(a) $4e^x, e^x, 5e^x$
(b) $\cos x, e^{ix}, 3\sin x$
(c) $7, x^2, 9x^4, e^{-x}$

Exercise 2.8
Are the following sets of functions linearly independent or dependent on the positive x-axis?
(a) $x, x+2, x+5$
(b) $\cos x, \cos 2x, \cos 3x$
(c) $\sin^2 x, \cos^2 x, \sin 2x$
(d) $x, (x-1)^2, (x+1)^2$
(e) $\sinh^2 x, \cosh^2 x, 1$

Exercise 2.9
Are the following sets of vectors linearly independent or dependent over the complex field?
(a) $(2, -3, 0), (0, 0, 1), (2i, i, -i)$
(b) $(0, 4, 0), (i, -3i, i), (2, 0, 1)$
(c) $(i, 1, 2), (3, i, -1), (-i, 3i, 5i)$

Exercise 2.10
Are the following sets of vectors (in the three-dimensional Euclidean space) linearly independent or dependent?

(a) $(4,5,6), (1,2,3), (7,8,9)$
(b) $(1,0,0), (0,-5,0), (0,0,\sqrt{7})$
(c) $(5,4,1), (2,0,-2), (0,6,-1)$

Exercise 2.11
Show that if $\hat{A}$ is a projection operator, the operator $1 - \hat{A}$ is also a projection operator.

Exercise 2.12
Show that $|\psi\rangle\langle\psi| / \langle\psi|\psi\rangle$ is a projection operator, regardless of whether $|\psi\rangle$ is normalized or not.

Exercise 2.13
In the following expressions, where $\hat{A}$ is an operator, specify the nature of each expression (i.e., specify whether it is an operator, a bra, or a ket); then find its Hermitian conjugate.

(a) $\langle\phi|\hat{A}|\psi\rangle\langle\psi|$
(b) $\hat{A}|\psi\rangle\langle\phi|$
(c) $\langle\phi|\hat{A}|\psi\rangle|\psi\rangle\langle\phi|\hat{A}$
(d) $\langle\psi|\hat{A}|\phi\rangle|\phi\rangle + i\hat{A}|\psi\rangle$
(e) $\left(|\phi\rangle\langle\phi|\hat{A}\right) - i\left(\hat{A}|\psi\rangle\langle\psi|\right)$

Exercise 2.14
Consider a two-dimensional space where a Hermitian operator $\hat{A}$ is defined by $\hat{A}|\phi_1\rangle = |\phi_1\rangle$ and $\hat{A}|\phi_2\rangle = -|\phi_2\rangle$; $|\phi_1\rangle$ and $|\phi_2\rangle$ are orthonormal.

(a) Do the states $|\phi_1\rangle$ and $|\phi_2\rangle$ form a basis?
(b) Consider the operator $\hat{B} = |\phi_1\rangle\langle\phi_2|$. Is $\hat{B}$ Hermitian? Show that $\hat{B}^2 = 0$.
(c) Show that the products $\hat{B}\hat{B}^\dagger$ and $\hat{B}^\dagger\hat{B}$ are projection operators.
(d) Show that the operator $\hat{B}\hat{B}^\dagger - \hat{B}^\dagger\hat{B}$ is unitary.
(e) Consider $\hat{C} = \hat{B}\hat{B}^\dagger + \hat{B}^\dagger\hat{B}$. Show that $\hat{C}|\phi_1\rangle = |\phi_1\rangle$ and $\hat{C}|\phi_2\rangle = |\phi_2\rangle$.

Exercise 2.15
Prove the following two relations:

(a) $e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}}e^{[\hat{A},\hat{B}]/2}$,
(b) $e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A},\hat{B}] + \frac{1}{2!}[\hat{A},[\hat{A},\hat{B}]] + \frac{1}{3!}[\hat{A},[\hat{A},[\hat{A},\hat{B}]]] + \cdots$.

Hint: To prove the first relation, you may consider defining an operator function $\hat{F}(t) = e^{\hat{A}t}e^{\hat{B}t}$, where t is a parameter, $\hat{A}$ and $\hat{B}$ are t-independent operators, and then make use of $[\hat{A}, G(\hat{B})] = [\hat{A},\hat{B}]dG(\hat{B})/d\hat{B}$, where $G(\hat{B})$ is a function depending on the operator $\hat{B}$.

Exercise 2.16
(a) Verify that the matrix

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

is unitary.
(b) Find its eigenvalues and the corresponding normalized eigenvectors.

Exercise 2.17
Consider the following three matrices:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(a) Calculate the commutators $[A, B]$, $[B, C]$, and $[C, A]$.
(b) Show that $A^2 + B^2 + 2C^2 = 4I$, where I is the unity matrix.
(c) Verify that $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB)$.

Exercise 2.18
Consider the following two matrices:

$$A = \begin{pmatrix} 3 & i & 1 \\ -1 & -i & 2 \\ 4 & 3i & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2i & 5 & -3 \\ -i & 3 & 0 \\ 7i & 1 & i \end{pmatrix}.$$

Verify the following relations:
(a) $\det(AB) = \det(A)\det(B)$,
(b) $\det(A^T) = \det(A)$,
(c) $\det(A^\dagger) = (\det(A))^*$, and
(d) $\det(A^*) = (\det(A))^*$.

Exercise 2.19
Consider the matrix

$$A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

(a) Find the eigenvalues and the normalized eigenvectors for the matrix A.
(b) Do these eigenvectors form a basis (i.e., is this basis complete and orthonormal)?
(c) Consider the matrix U which is formed from the normalized eigenvectors of A. Verify that U is unitary and that it satisfies

$$U^\dagger A U = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

where λ_1 and λ_2 are the eigenvalues of A.
(d) Show that $e^{xA} = \cosh x + A \sinh x$.

Exercise 2.20
Using the bra-ket algebra, show that $\mathrm{Tr}(\hat{A}\hat{B}\hat{C}) = \mathrm{Tr}(\hat{C}\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{C}\hat{A})$ where $\hat{A}, \hat{B}, \hat{C}$ are operators.

Exercise 2.21
For any two kets $|\psi\rangle$ and $|\phi\rangle$ that have finite norm, show that $\mathrm{Tr}\left(|\psi\rangle\langle\phi|\right) = \langle\phi|\psi\rangle$.

Exercise 2.22

Consider the matrix $A = \begin{pmatrix} 0 & 0 & -1+i \\ 0 & 3 & 0 \\ -1-i & 0 & 0 \end{pmatrix}$.

(a) Find the eigenvalues and normalized eigenvectors of A. Denote the eigenvectors of A by $|a_1\rangle, |a_2\rangle, |a_3\rangle$. Any degenerate eigenvalues?

(b) Show that the eigenvectors $|a_1\rangle, |a_2\rangle, |a_3\rangle$ form an orthonormal and complete basis, i.e., show that $\sum_{j=1}^{3} |a_j\rangle\langle a_j| = I$, where I is the 3×3 unit matrix, and that $\langle a_j | a_k\rangle = \delta_{jk}$.

(c) Find the matrix corresponding to the operator obtained from the ket-bra product of the first eigenvector $P = |a_1\rangle\langle a_1|$. Is P a projection operator?

Exercise 2.23

In a three-dimensional vector space, consider the operator whose matrix, in an orthonormal basis $\{|1\rangle, |2\rangle, |3\rangle\}$, is

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

(a) Is A Hermitian? Calculate its eigenvalues and the corresponding normalized eigenvectors. Verify that the eigenvectors corresponding to the two nondegenerate eigenvalues are orthonormal.

(b) Calculate the matrices representing the projection operators for the two nondegenerate eigenvectors found in part (a).

Exercise 2.24

Consider two operators $\hat{A}$ and $\hat{B}$ whose matrices are

$$A = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 4 \end{pmatrix}.$$

(a) Are $\hat{A}$ and $\hat{B}$ Hermitian?

(b) Do $\hat{A}$ and $\hat{B}$ commute?

(c) Find the eigenvalues and eigenvectors of $\hat{A}$ and $\hat{B}$.

(d) Are the eigenvectors of each operator orthonormal?

(e) Verify that $\hat{U}^\dagger \hat{B} \hat{U}$ is diagonal, $\hat{U}$ being the matrix of the normalized eigenvectors of $\hat{B}$.

(f) Verify that $\hat{U}^{-1} = \hat{U}^\dagger$.

Exercise 2.25

Consider an operator $\hat{A}$ so that $[\hat{A}, \hat{A}^\dagger] = 1$.

(a) Evaluate the commutators $[\hat{A}^\dagger \hat{A}, \hat{A}]$ and $[\hat{A}^\dagger \hat{A}, \hat{A}^\dagger]$.

(b) If the actions of $\hat{A}$ and $\hat{A}^\dagger$ on the states $\{|a\rangle\}$ are given by $\hat{A}\,|a\rangle = \sqrt{a}\,|a-1\rangle$ and $\hat{A}^\dagger\,|a\rangle = \sqrt{a+1}\,|a+1\rangle$ and if $\langle a' | a\rangle = \delta_{a'a}$, calculate $\langle a | \hat{A} | a+1\rangle$, $\langle a+1 | \hat{A}^\dagger | a\rangle$ and $\langle a | \hat{A}^\dagger \hat{A} | a\rangle$ and $\langle a | \hat{A}\hat{A}^\dagger | a\rangle$.

(c) Calculate $\langle a | (\hat{A} + \hat{A}^\dagger)^2 | a\rangle$ and $\langle a | (\hat{A} - \hat{A}^\dagger)^2 | a\rangle$.

Exercise 2.26
Consider a 4×4 matrix

$$A = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(a) Find the matrices of $A^\dagger$, $N = A^\dagger A$, $H = N + \frac{1}{2}I$ (where I is the unit matrix), $B = A + A^\dagger$, and $C = i(A - A^\dagger)$.

(b) Find the matrices corresponding to the commutators $[A^\dagger, A]$, $[B, C]$, $[N, B]$, and $[N, C]$.

(c) Find the matrices corresponding to B^2, C^2, $[N, B^2 + C^2]$, $[H, A^\dagger]$, $[H, A]$, and $[H, N]$.

(d) Verify that $\det(ABC) = \det(A)\det(B)\det(C)$ and $\det(C^\dagger) = (\det(C))^*$.

Exercise 2.27
If $\hat{A}$ and $\hat{B}$ commute, and if $|\psi_1\rangle$ and $|\psi_2\rangle$ are two eigenvectors of $\hat{A}$ with different eigenvalues ($\hat{A}$ is Hermitian), show that

(a) $\langle\psi_1 | \hat{B} | \psi_2\rangle$ is zero and

(b) $\hat{B}|\psi_1\rangle$ is also an eigenvector to $\hat{A}$ with the same eigenvalue as $|\psi_1\rangle$; i.e., if $\hat{A}|\psi_1\rangle = a_1|\psi_1\rangle$, show that $\hat{A}(\hat{B}|\psi_1\rangle) = a_1\hat{B}|\psi_1\rangle$.

Exercise 2.28
Let A and B be two $n \times n$ matrices. Assuming that B^{-1} exists, show that $[A, B^{-1}] = -B^{-1}[A, B]B^{-1}$.

Exercise 2.29
Consider a physical system whose Hamiltonian H and an operator A are given, in a three-dimensional space, by the matrices

$$H = \hbar\omega \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad A = a \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

(a) Are H and A Hermitian?

(b) Show that H and A commute. Give a basis of eigenvectors common to H and A.

Exercise 2.30
(a) Using $[\hat{X}, \hat{P}] = i\hbar$, show that $[\hat{X}^2, \hat{P}] = 2i\hbar\hat{X}$ and $[\hat{X}, \hat{P}^2] = 2i\hbar\hat{P}$.

(b) Show that $[\hat{X}^2, \hat{P}^2] = 2i\hbar(i\hbar + 2\hat{P}\hat{X})$.

(c) Calculate the commutator $[\hat{X}^2, \hat{P}^3]$.

Exercise 2.31
Discuss the hermiticity of the commutators $[\hat{X}, \hat{P}]$, $[\hat{X}^2, \hat{P}]$ and $[\hat{X}, \hat{P}^2]$.

Exercise 2.32
(a) Evaluate the commutator $[\hat{X}^2, d/dx]$ by operating it on a wave function.

(b) Using $[\hat{X}, \hat{P}] = i\hbar$, evaluate the commutator $[\hat{X}\hat{P}^2, \hat{P}\hat{X}^2]$ in terms of a linear combination of $\hat{X}^2\hat{P}^2$ and $\hat{X}\hat{P}$.

Exercise 2.33
Show that $[\hat{X}, \hat{P}^n] = i\hbar n\hat{P}^{n-1}$.

Exercise 2.34
Evaluate the commutators $[e^{i\hat{X}}, \hat{P}]$, $[e^{i\hat{X}^2}, \hat{P}]$, and $[e^{i\hat{X}}, \hat{P}^2]$.

Exercise 2.35
Consider the matrix

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

(a) Find the eigenvalues and the normalized eigenvectors of A.
(b) Do these eigenvectors form a basis (i.e., is this basis complete and orthonormal)?
(c) Consider the matrix U which is formed from the normalized eigenvectors of A. Verify that U is unitary and that it satisfies the relation

$$U^\dagger AU = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

where λ_1, λ_2, and λ_3 are the eigenvalues of A.
(d) Show that $e^{xA} = \cosh x + A \sinh x$.
Hint: $\cosh x = \sum_{n=0}^{\infty} x^{2n}/(2n)!$ and $\sinh x = \sum_{n=0}^{\infty} x^{2n+1}/(2n+1)!$.

Exercise 2.36
(a) If $[\hat{A}, \hat{B}] = c$, where c is a number, prove the following two relations: $e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + c$ and $e^{\hat{A}+\hat{B}} = e^{\hat{A}}e^{\hat{B}}e^{-c/2}$.
(b) Now if $[\hat{A}, \hat{B}] = c\hat{B}$, where c is again a number, show that $e^{\hat{A}}\hat{B}e^{-\hat{A}} = e^{c}\hat{B}$.

Exercise 2.37
Consider the matrix

$$A = \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & 3 \end{pmatrix}.$$

(a) Find the eigenvalues of A and their corresponding eigenvectors.
(b) Consider the basis which is constructed from the three eigenvectors of A. Using matrix algebra, verify that this basis is both orthonormal and complete.

Exercise 2.38
(a) Specify the condition that must be satisfied by a matrix A so that it is both unitary and Hermitian.
(b) Consider the three matrices

$$M_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Calculate the inverse of each matrix. Do they satisfy the condition derived in (a)?

Exercise 2.39
Consider the two matrices

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}, \qquad B = \frac{1}{2}\begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}.$$

(a) Are these matrices Hermitian?
(b) Calculate the inverses of these matrices.
(c) Are these matrices unitary?
(d) Verify that the determinants of A and B are of the form $e^{i\theta}$. Find the corresponding values of θ.

Exercise 2.40
Show that the transformation matrix representing a 90° counterclockwise rotation about the z-axis of the basis vectors $(\hat{\imath}, \hat{\jmath}, \hat{k})$ is given by

$$U = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Exercise 2.41
Show that the transformation matrix representing a 90° clockwise rotation about the y-axis of the basis vectors $(\hat{\imath}, \hat{\jmath}, \hat{k})$ is given by

$$U = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Exercise 2.42
Show that the operator $(\hat{X}\hat{P} + \hat{P}\hat{X})^2$ is equal to $(\hat{X}^2\hat{P}^2 + \hat{P}^2\hat{X}^2)$ plus a term of the order of $\hbar^2$.

Exercise 2.43
Consider the two matrices $A = \begin{pmatrix} 4 & i & 7 \\ 1 & 0 & 1 \\ 0 & 1 & -i \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & i & 0 \\ -i & 0 & i \end{pmatrix}$. Calculate the products $B^{-1}A$ and AB^{-1}. Are they equal? What is the significance of this result?

Exercise 2.44
Use the relations listed in Appendix A to evaluate the following integrals involving Dirac's delta function:
(a) $\int_0^{\pi} \sin(3x)\cos^2(4x)\delta(x - \pi/2)\,dx$.
(b) $\int_{-2}^{2} e^{7x+2}\delta(5x)\,dx$.
(c) $\int_{-2\pi}^{2\pi} \sin(\theta/2)\delta''(\theta + \pi)\,d\theta$.
(d) $\int_0^{2\pi} \cos^2\theta\,\delta[(\theta - \pi)/4]\,d\theta$.

Exercise 2.45
Use the relations listed in Appendix A to evaluate the following expressions:
(a) $\int_0^5 (3x^2 + 2)\delta(x - 1)\,dx$.
(b) $(2x^5 - 4x^3 + 1)\delta(x + 2)$.
(c) $\int_0^{\infty} (5x^3 - 7x^2 - 3)\delta(x^2 - 4)\,dx$.

Exercise 2.46
Use the relations listed in Appendix A to evaluate the following expressions:

(a) $\int_3^7 e^{6x-2}\delta(-4x)\,dx$.

(b) $\cos(2\theta)\sin(\theta)\delta(\theta^2 - \pi^2/4)$.

(c) $\int_{-1}^{1} e^{5x-1}\delta'''(x)\,dx$.

Exercise 2.47
If the position and momentum operators are denoted by $\hat{\vec{R}}$ and $\hat{\vec{P}}$, respectively, show that $\hat{\mathcal{P}}^\dagger \hat{\vec{R}}^n \hat{\mathcal{P}} = (-1)^n \hat{\vec{R}}^n$ and $\hat{\mathcal{P}}^\dagger \hat{\vec{P}}^n \hat{\mathcal{P}} = (-1)^n \hat{\vec{P}}^n$, where $\hat{\mathcal{P}}$ is the parity operator and n is an integer.

Exercise 2.48
Consider an operator

$$\begin{aligned}\hat{A} \;=\; & |\,\phi_1\rangle\langle\phi_1\,| + |\,\phi_2\rangle\langle\phi_2\,| + |\,\phi_3\rangle\langle\phi_3\,| - i\,|\,\phi_1\rangle\langle\phi_2\,| \\ & - |\,\phi_1\rangle\langle\phi_3\,| + i\,|\,\phi_2\rangle\langle\phi_1\,| - |\,\phi_3\rangle\langle\phi_1\,|,\end{aligned}$$

where $|\,\phi_1\rangle$, $|\,\phi_2\rangle$, and $|\,\phi_3\rangle$ form a complete and orthonormal basis.

(a) Is $\hat{A}$ Hermitian? Calculate $\hat{A}^2$; is it a projection operator?

(b) Find the 3×3 matrix representing $\hat{A}$ in the $|\,\phi_1\rangle$, $|\,\phi_2\rangle$, $|\,\phi_3\rangle$ basis.

(c) Find the eigenvalues and the eigenvectors of the matrix.

Exercise 2.49
The Hamiltonian of a two-state system is given by

$$\hat{H} = E\big(|\,\phi_1\rangle\langle\phi_1\,| - |\,\phi_2\rangle\langle\phi_2\,| - i\,|\,\phi_1\rangle\langle\phi_2\,| + i\,|\,\phi_2\rangle\langle\phi_1\,|\big),$$

where $|\,\phi_1\rangle$, $|\,\phi_2\rangle$ form a complete and orthonormal basis; E is a real constant having the dimensions of energy.

(a) Is $\hat{H}$ Hermitian? Calculate the trace of $\hat{H}$.

(b) Find the matrix representing $\hat{H}$ in the $|\,\phi_1\rangle$, $|\,\phi_2\rangle$ basis and calculate the eigenvalues and the eigenvectors of the matrix. Calculate the trace of the matrix and compare it with the result you obtained in (a).

(c) Calculate $[\hat{H}, |\,\phi_1\rangle\langle\phi_1\,|]$, $[\hat{H}, |\,\phi_2\rangle\langle\phi_2\,|]$, and $[\hat{H}, |\,\phi_1\rangle\langle\phi_2\,|]$.

Exercise 2.50
Consider a particle which is confined to move along the positive x-axis and whose Hamiltonian is $\hat{H} = \mathcal{E}d^2/dx^2$, where $\mathcal{E}$ is a positive real constant having the dimensions of energy.

(a) Find the wave function that corresponds to an energy eigenvalue of $9\mathcal{E}$ (make sure that the function you find is finite everywhere along the positive x-axis and is square integrable). Normalize this wave function.

(b) Calculate the probability of finding the particle in the region $0 \le x \le 15$.

(c) Is the wave function derived in (a) an eigenfunction of the operator $\hat{A} = d/dx - 7$?

(d) Calculate the commutator $[\hat{H}, \hat{A}]$.

Exercise 2.51
Consider the wave functions:

$$\psi(x, y) = \sin 2x \cos 5x, \qquad \phi(x, y) = e^{-2(x^2+y^2)}, \qquad \chi(x, y) = e^{-i(x+y)}.$$

(a) Verify if any of the wave functions is an eigenfunction of $\hat{A} = \partial/\partial x + \partial/\partial y$.
(b) Find out if any of the wave functions is an eigenfunction of $\hat{B} = \partial^2/\partial x^2 + \partial^2/\partial y^2 + 1$.
(c) Calculate the actions of $\hat{A}\hat{B}$ and $\hat{B}\hat{A}$ on each one of the wave functions and infer $[\hat{A}, \hat{B}]$.

Exercise 2.52
(a) Is the state $\psi(\theta, \phi) = e^{-3i\phi} \cos\theta$ an eigenfunction of $\hat{A}_\phi = \partial/\partial\phi$ or of $\hat{B}_\theta = \partial/\partial\theta$?
(b) Are $\hat{A}_\phi$ and $\hat{B}_\theta$ Hermitian?
(c) Evaluate the expressions $\langle\psi \mid \hat{A}_\phi \mid \psi\rangle$ and $\langle\psi \mid \hat{B}_\theta \mid \psi\rangle$.
(d) Find the commutator $[\hat{A}_\phi, \hat{B}_\theta]$.

Exercise 2.53
Consider an operator $\hat{A} = (\hat{X}d/dx + 2)$.
(a) Find the eigenfunction of $\hat{A}$ corresponding to a zero eigenvalue. Is this function normalizable?
(b) Is the operator $\hat{A}$ Hermitian?
(c) Calculate $[\hat{A}, \hat{X}]$, $[\hat{A}, d/dx]$, $[\hat{A}, d^2/dx^2]$, $[\hat{X}, [\hat{A}, \hat{X}]]$, and $[d/dx, [\hat{A}, d/dx]]$.

Exercise 2.54
If $\hat{A}$ and $\hat{B}$ are two Hermitian operators, find their respective eigenvalues such that $\hat{A}^2 = 2\hat{I}$ and $\hat{B}^4 = \hat{I}$, where $\hat{I}$ is the unit operator.

Exercise 2.55
Consider the Hilbert space of two-variable complex functions $\psi(x, y)$. A permutation operator is defined by its action on $\psi(x, y)$ as follows: $\hat{\pi}\psi(x, y) = \psi(y, x)$.
(a) Verify that the operator $\hat{\pi}$ is linear and Hermitian.
(b) Show that $\hat{\pi}^2 = \hat{I}$. Find the eigenvalues and show that the eigenfunctions of $\hat{\pi}$ are given by

$$\psi_+(x, y) = \frac{1}{2}[\psi(x, y) + \psi(y, x)] \quad \text{and} \quad \psi_-(x, y) = \frac{1}{2}[\psi(x, y) - \psi(y, x)].$$

Chapter 3

Postulates of Quantum Mechanics

3.1 Introduction

The formalism of quantum mechanics is based on a number of postulates. These postulates are in turn based on a wide range of experimental observations; the underlying physical ideas of these experimental observations have been briefly mentioned in Chapter 1. In this chapter we present a formal discussion of these postulates, and how they can be used to extract quantitative information about microphysical systems.

These postulates cannot be derived; they result from experiment. They represent the minimal set of assumptions needed to develop the theory of quantum mechanics. But how does one find out about the validity of these postulates? Their validity cannot be determined directly; only an indirect inferential statement is possible. For this, one has to turn to the theory built upon these postulates: if the theory works, the postulates will be valid; otherwise they will make no sense. Quantum theory not only works, but works extremely well, and this represents its experimental justification. It has a very penetrating qualitative as well as quantitative prediction power; this prediction power has been verified by a rich collection of experiments. So the accurate prediction power of quantum theory gives irrefutable evidence to the validity of the postulates upon which the theory is built.

3.2 The Basic Postulates of Quantum Mechanics

According to classical mechanics, the state of a particle is specified, at any time t, by two fundamental dynamical variables: the position $\vec{r}(t)$ and the momentum $\vec{p}(t)$. Any other physical quantity, relevant to the system, can be calculated in terms of these two dynamical variables. In addition, knowing these variables at a time t, we can predict, using for instance Hamilton's equations $dx/dt = \partial H/\partial p$ and $dp/dt = -\partial H/\partial x$, the values of these variables at any later time t'.

The quantum mechanical counterparts to these ideas are specified by postulates, which enable us to understand:

- how a quantum state is described mathematically at a given time t,
- how to calculate the various physical quantities from this quantum state, and

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

- knowing the system's state at a time t, how to find the state at any later time t'; that is, how to describe the time evolution of a system.

The answers to these questions are provided by the following set of five postulates.

Postulate 1: State of a system
The state of any physical system is specified, at each time t, by a state vector $|\psi(t)\rangle$ in a Hilbert space $\mathcal{H}$; $|\psi(t)\rangle$ contains (and serves as the basis to extract) all the needed information about the system. Any superposition of state vectors is also a state vector.

Postulate 2: Observables and operators
To every physically measurable quantity A, called an observable or dynamical variable, there corresponds a linear Hermitian operator $\hat{A}$.

Postulate 3: Measurements and eigenvalues of operators
The measurement of an observable A may be represented formally by the action of $\hat{A}$ on a state vector $|\psi(t)\rangle$. The only possible result of such a measurement is one of the eigenvalues of the operator $\hat{A}$.
Remark
If the spectrum of $\hat{A}$ is discrete and if all of its eigenvalues $(a_1, a_2, a_3, \cdots)$ are not degenerate, every eigenvalue a_n has a corresponding and unique eigenvector $|\phi_n\rangle$:

$$\hat{A}|\phi_n\rangle = a_n|\phi_n\rangle. \tag{3.1}$$

If $\hat{A}$ is Hermitian, the set of $|\phi_n\rangle$ will form a *complete basis*, $\sum_n |\phi_n\rangle\langle\phi_n| = \hat{I}$. In order for $\hat{A}$ to be associated with an observable, it must possible to expand any state vector $|\psi(t)\rangle$ in terms of the eigenvectors of $\hat{A}$:

$$|\psi(t)\rangle = \sum_n c_n|\phi_n\rangle, \tag{3.2}$$

where c_n is the component of $|\psi(t)\rangle$ when projected[1] onto the eigenvectors $|\phi_n\rangle$: $c_n = \langle\phi_n|\psi(t)\rangle$.

Postulate 4: Probabilistic outcome of measurements

- **Postulate 4 for discrete spectra**: *When measuring an observable A of a system in a state $|\psi\rangle$, the probability of obtaining one of the nondegenerate eigenvalues a_n of the corresponding operator $\hat{A}$ is given by*:

$$\boxed{P_n(a_n) = \frac{|\langle\phi_n|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \frac{|c_n|^2}{\langle\psi|\psi\rangle},} \tag{3.3}$$

  where $|\phi_n\rangle$ is the eigenstate of $\hat{A}$ with eigenvalue a_n. We can verify that total probability is equal to 1. Using (3.3), we can write:

$$\sum_n P_n(a_n) = \frac{\sum_n |\langle\phi_n|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \frac{\langle\psi\left|\left(\sum_n |\phi_n\rangle\langle\phi_n\right)\right|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\langle\psi|\psi\rangle}{\langle\psi|\psi\rangle} = 1. \tag{3.4}$$

[1]Since the eigenvectors of $\hat{A}$ form a complete basis, we have: $|\psi(t)\rangle = \sum_n |\phi_n\rangle\langle\phi_n|\psi(t)\rangle = \sum_n c_n|\phi_n\rangle$.

Now, if the eigenvalue a_n is m-degenerate, P_n becomes

$$P_n(a_n) = \frac{\sum_{j=1}^{m} \left|\langle \phi_n^j | \psi \rangle\right|^2}{\langle \psi | \psi \rangle} = \frac{\sum_{j=1}^{m} \left|c_n^{(j)}\right|^2}{\langle \psi | \psi \rangle}. \tag{3.5}$$

The act of measurement changes the state of the system from $|\psi\rangle$ to $|\phi_n\rangle$. If the system is already in an eigenstate $|\phi_n\rangle$ of $\hat{A}$, a measurement of A yields with certainty the corresponding eigenvalue a_n, since $\hat{A}|\phi_n\rangle = a_n|\phi_n\rangle$.

- **Continuous spectra:** If the spectrum of $\hat{A}$ is continuous and, for simplicity, if we assume that it is not degenerate, for every eigenvalue k (which is now a continuous number) corresponds a unique eigenstate $|\chi_k\rangle$:

$$\hat{A}\,|\chi_k\rangle = k\,|\chi_k\rangle, \tag{3.6}$$

If $\hat{A}$ is Hermitian, the eigenvectors $|\chi_k\rangle$ will form a *complete basis*, $\int_{-\infty}^{+\infty} dk \;|\chi_k\rangle\langle\chi_k| = \hat{I}$; the state vector $|\psi(t)\rangle$ can be expanded as follows:

$$|\psi(t)\rangle = \int dk\, c(k)\,|\chi_k\rangle, \tag{3.7}$$

where $c(k)$ is the component of $|\psi(t)\rangle$ when projected[2] onto the eigenvector $|\chi_k\rangle$: $c(k) = \langle\chi_k|\psi(t)\rangle$.

The relation (3.3), which is valid for discrete spectra, can be extended easily to continuous spectra as follows: **Postulate 4 for continuous spectra:** *the probability density that a measurement of $\hat{A}$ yields a value between k and $k + dk$ on a system which is initially in a state $|\psi\rangle$ is given by:*

$$\boxed{\frac{dP(k)}{dk} = \frac{\left|\langle\chi_k|\psi(t)\rangle\right|^2}{\langle\psi|\psi\rangle} = \frac{|c(k)|^2}{\langle\psi|\psi\rangle};} \tag{3.8}$$

for instance, the probability density for finding a particle between x and $x + dx$ is given by $dP(x)/dx = |\psi(x)|^2/\langle\psi|\psi\rangle$.

Using (3.8), we can verify that the total probability is equal to 1:

$$\int_{-\infty}^{+\infty} dP(k) = \frac{1}{\langle\psi|\psi\rangle}\int_{-\infty}^{+\infty} dk\;\left|\langle\chi_k|\psi(t)\rangle\right|^2 = \frac{1}{\langle\psi|\psi\rangle}\langle\psi\left|\left(\int_{-\infty}^{+\infty} dk\;|\chi_k\rangle\langle\chi_k|\right)\right|\psi\rangle = \frac{\langle\psi|\psi\rangle}{\langle\psi|\psi\rangle} = 1. \tag{3.9}$$

Postulate 5: Time evolution of a system

The time evolution of the state vector $|\psi(t)\rangle$ of a system is governed by the time-dependent Schrödinger equation

$$\boxed{i\hbar\frac{\partial|\psi(t)\rangle}{\partial t} = \hat{H}|\psi(t)\rangle,} \tag{3.10}$$

[2]Since the eigenvectors of $\hat{A}$ form a complete basis, we have: $|\psi(t)\rangle = \int dk|\chi_k\rangle\langle\chi_k|\psi(t)\rangle = \int dk\, c(k)\,|\chi_k\rangle$.

where $\hat{H}$ is the Hamiltonian operator corresponding to the total energy of the system.

Remarks
These postulates fall into two categories:

- The first four describe the system at a given time.
- The fifth shows how this description evolves in time.

In the rest of this chapter we are going to consider the physical implications of each one of the four postulates. Namely, we shall look at the state of a quantum system and its interpretation, the physical observables, measurements in quantum mechanics, and finally the time evolution of quantum systems.

3.3 The State of a System

To describe a system in quantum mechanics, we use a mathematical entity (a complex function) belonging to a Hilbert space, the state vector $|\psi(t)\rangle$, which contains all the information we need to know about the system and from which all needed physical quantities can be computed. As discussed in Chapter 2, the state vector $|\psi(t)\rangle$ may be represented in two ways:

- A wave function $\psi(\vec{r},t)$ in the position space: $\psi(\vec{r},t) = \langle\vec{r}|\psi(t)\rangle$.
- A momentum wave function $\Psi(\vec{p},t)$ in the momentum space: $\Psi(\vec{p},t) = \langle\vec{p}|\psi(t)\rangle$.

So, for instance, to describe the state of a one-dimensional particle in quantum mechanics we use a complex function $\psi(x,t)$ instead of two real real numbers (x,p) in classical physics.

The wave functions to be used are only those that correspond to physical systems. What are the mathematical requirements that a wave function must satisfy to represent a physical system? Wave functions $\psi(x)$ that are physically acceptable must, along with their first derivatives $d\psi(x)/dx$, be *finite, continuous,* and *single-valued everywhere.* As will be discussed in Chapter 4, we will examine the underlying physics behind the continuity conditions of $\psi(x)$ and $d\psi(x)/dx$ (we will see that $\psi(x)$ and $d\psi(x)/dx$ must be be continuous because the probability density and the linear momentum are continuous functions of x).

3.3.1 Probability Density

What about the physical meaning of a wave function? Only the square of its norm, $|\psi(\vec{r},t)|^2$, has meaning. According to Born's probabilistic interpretation, the square of the norm of $\psi(\vec{r},t)$,

$$P(\vec{r},t) = |\psi(\vec{r},t)|^2, \tag{3.11}$$

represents a position probability density; that is, the quantity $|\psi(\vec{r},t)|^2 d^3r$ represents the probability of finding the particle at time t in a volume element d^3r located between $\vec{r}$ and $\vec{r}+d\vec{r}$. Therefore, the total probability of finding the system somewhere in space is equal to 1:

$$\int |\psi(\vec{r},t)|^2 d^3r = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} |\psi(\vec{r},t)|^2 dz = 1. \tag{3.12}$$

A wave function $\psi(\vec{r},t)$ satisfying this relation is said to be *normalized*. We may mention that $\psi(\vec{r})$ has the physical dimensions of $1/\sqrt{L^3}$, where L is a length. Hence, the physical dimensions of $|\psi(\vec{r})|^2$ is $1/L^3$: $\left[|\psi(\vec{r})|^2\right] = 1/L^3$.

Note that the wave functions $\psi(\vec{r},t)$ and $e^{i\alpha}\psi(\vec{r},t)$, where α is a real number, represent the same state.

Example 3.1 (Physical and unphysical wave functions)
Which among the following functions represent physically acceptable wave functions: $f(x) = 3\sin \pi x$, $g(x) = 4 - |x|$, $h^2(x) = 5x$, and $e(x) = x^2$.

Solution
Among these functions only $f(x) = 3\sin \pi x$ represents a physically acceptable wave function, since $f(x)$ and its derivative are finite, continuous, single-valued everywhere, and integrable.

The other functions cannot be wave functions, since $g(x) = 4 - |x|$ is not continuous, not finite, and not square integrable; $h^2(x) = 5x$ is neither finite nor square integrable; and $e(x) = x^2$ is neither finite nor square integrable.

3.3.2 The Superposition Principle

The state of a system does not have to be represented by a *single* wave function; it can be represented by a *superposition* of two or more wave functions. An example from the macroscopic world is a vibrating string; its state can be represented by a single wave or by the superposition (linear combination) of many waves.

If $\phi_1(\vec{r},t)$ and $\phi_2(\vec{r},t)$ separately satisfy the Schrödinger equation, then the wave function $\psi(\vec{r},t) = \alpha_1\phi_1(\vec{r},t) + \alpha_2\phi_2(\vec{r},t)$ also satisfies the Schrödinger equation, where α_1 and α_2 are complex numbers. The Schrödinger equation is a linear equation. So in general, according to the superposition principle, the linear superposition of many wave functions (which describe the various permissible physical states of a system) gives a new wave function which represents a possible physical state of the system:

$$|\psi\rangle = \sum_i \alpha_i |\phi_i\rangle, \tag{3.13}$$

where the α_i are complex numbers. The quantity

$$P = \left|\sum_i \alpha_i |\phi_i\rangle\right|^2, \tag{3.14}$$

represents the probability for this superposition. If the states $|\phi_i\rangle$ are mutually *orthonormal*, the probability will be equal to the sum of the individual probabilities:

$$P = \left|\sum_i \alpha_i |\phi_i\rangle\right|^2 = \sum_i |\alpha_i|^2 = P_1 + P_2 + P_3 + \cdots, \tag{3.15}$$

where $P_i = |\alpha_i|^2$; P_i is the probability of finding the system in the state $|\phi_i\rangle$.

Example 3.2
Consider a system whose state is given in terms of an orthonormal set of three vectors: $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$ as

$$|\psi\rangle = \frac{\sqrt{3}}{3}|\phi_1\rangle + \frac{2}{3}|\phi_2\rangle + \frac{\sqrt{2}}{3}|\phi_3\rangle.$$

(a) Verify that $|\psi\rangle$ is normalized. Then, calculate the probability of finding the system in any one of the states $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$. Verify that the total probability is equal to one.

(b) Consider now an ensemble of 810 identical systems, each one of them in the state $|\psi\rangle$. If measurements are done on all of them, how many systems will be found in each of the states $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$?

Solution

(a) Using the orthonormality condition $\langle\phi_j|\phi_k\rangle = \delta_{jk}$ where $j, k = 1, 2, 3$, we can verify that $|\psi\rangle$ is normalized:

$$\langle\psi|\psi\rangle = \frac{1}{3}\langle\phi_1|\phi_1\rangle + \frac{4}{9}\langle\phi_2|\phi_2\rangle + \frac{2}{9}\langle\phi_3|\phi_3\rangle = \frac{1}{3} + \frac{4}{9} + \frac{2}{9} = 1. \tag{3.16}$$

Since $|\psi\rangle$ is normalized, the probability of finding the system in $|\phi_1\rangle$ is given by

$$P_1 = |\langle\phi_1|\psi\rangle|^2 = \left|\frac{\sqrt{3}}{3}\langle\phi_1|\phi_1\rangle + \frac{2}{3}\langle\phi_1|\phi_2\rangle + \frac{\sqrt{2}}{3}\langle\phi_1|\phi_3\rangle\right|^2 = \frac{1}{3}, \tag{3.17}$$

since $\langle\phi_1|\phi_1\rangle = 1$ and $\langle\phi_1|\phi_2\rangle = \langle\phi_1|\phi_3\rangle = 0$.

Similarly, from the relations $\langle\phi_2|\phi_2\rangle = 1$ and $\langle\phi_2|\phi_1\rangle = \langle\phi_2|\phi_3\rangle = 0$, we obtain the probability of finding the system in $|\phi_2\rangle$:

$$P_2 = |\langle\phi_2|\psi\rangle|^2 = \left|\frac{2}{3}\langle\phi_2|\phi_2\rangle\right|^2 = \frac{4}{9}. \tag{3.18}$$

As for $\langle\phi_3|\phi_3\rangle = 1$ and $\langle\phi_3|\phi_1\rangle = \langle\phi_3|\phi_2\rangle = 0$, they lead to the probability of finding the system in $|\phi_3\rangle$:

$$P_3 = |\langle\phi_3|\psi\rangle|^2 = \left|\frac{\sqrt{2}}{3}\langle\phi_3|\phi_3\rangle\right|^2 = \frac{2}{9}. \tag{3.19}$$

As expected, the total probability is equal to one:

$$P = P_1 + P_2 + P_3 = \frac{1}{3} + \frac{4}{9} + \frac{2}{9} = 1. \tag{3.20}$$

(b) The number of systems that will be found in the state $|\phi_1\rangle$ is

$$N_1 = 810 \times P_1 = \frac{810}{3} = 270. \tag{3.21}$$

Likewise, the number of systems that will be found in states $|\phi_2\rangle$ and $|\phi_3\rangle$ are given, respectively, by

$$N_2 = 810 \times P_2 = \frac{810 \times 4}{9} = 360, \qquad N_3 = 810 \times P_3 = \frac{810 \times 2}{9} = 180. \tag{3.22}$$

3.4 Observables and Operators

An observable is a dynamical variable that can be measured; the dynamical variables encountered most in classical mechanics are the position, linear momentum, angular momentum, and energy. How do we mathematically represent these and other variables in quantum mechanics?

According to the second postulate, a *Hermitian operator* is associated with every *physical observable*. In the preceding chapter, we have seen that the position representation of the linear momentum operator is given in one-dimensional space by $\hat{P} = -i\hbar\partial/\partial x$ and in three-dimensional space by $\hat{\vec{P}} = -i\hbar\vec{\nabla}$.

In general, any function, $f(\vec{r},\vec{p})$, which depends on the position and momentum variables, $\vec{r}$ and $\vec{p}$, can be "*quantized*" or made into a function of operators by replacing $\vec{r}$ and $\vec{p}$ with their corresponding operators:

$$f(\vec{r},\vec{p}) \quad \longrightarrow \quad F(\hat{\vec{R}},\hat{\vec{P}}) = f(\hat{\vec{R}}, -i\hbar\vec{\nabla}), \tag{3.23}$$

or $f(x,p) \rightarrow F(\hat{X},\ -i\hbar\partial/\partial x)$. For instance, the operator corresponding to the Hamiltonian

$$H = \frac{1}{2m}\vec{p}^{\,2} + V(\vec{r},t) \tag{3.24}$$

is given in the position representation by

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\hat{\vec{R}},t), \tag{3.25}$$

where ∇^2 is the Laplacian operator; it is given in Cartesian coordinates by: $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$.

Since the momentum operator $\hat{\vec{P}}$ is Hermitian, and if the potential $V(\hat{\vec{R}},t)$ is a real function, the Hamiltonian (3.24) is Hermitian. We saw in Chapter 2 that the eigenvalues of Hermitian operators are real. Hence, the spectrum of the Hamiltonian, which consists of the entire set of its eigenvalues, is real. This spectrum can be discrete, continuous, or a mixture of both. In the case of *bound* states, the Hamiltonian has a *discrete* spectrum of values and a *continuous* spectrum for *unbound* states. In general, an operator will have bound or unbound spectra in the same manner that the corresponding classical variable has bound or unbound orbits. As for $\hat{\vec{R}}$ and $\hat{\vec{P}}$, they have continuous spectra, since r and p may take a continuum of values.

According to Postulate 5, the total energy E for time-dependent systems is associated to the operator

$$\hat{H} = i\hbar\frac{\partial}{\partial t}. \tag{3.26}$$

This can be seen as follows. The wave function of a free particle of momentum $\vec{p}$ and total energy E is given by $\psi(\vec{r},t) = Ae^{i(\vec{p}\cdot\vec{r}-Et)/\hbar}$, where A is a constant. The time derivative of $\psi(\vec{r},t)$ yields

$$i\hbar\frac{\partial\psi(\vec{r},t)}{\partial t} = E\psi(\vec{r},t). \tag{3.27}$$

Table 3.1: Some observables and their corresponding operators.

Observable	Corresponding operator
$\vec{r}$	$\hat{\vec{R}}$
$\vec{p}$	$\hat{\vec{P}} = -i\hbar\vec{\nabla}$
$T = \frac{p^2}{2m}$	$\hat{T} = -\frac{\hbar^2}{2m}\nabla^2$
$E = \frac{p^2}{2m} + V(\vec{r},t)$	$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}(\hat{\vec{R}},t)$
$\vec{L} = \vec{r}\times\vec{p}$	$\hat{\vec{L}} = -i\hbar\vec{R}\times\vec{\nabla}$

Let us look at the eigenfunctions and eigenvalues of the momentum operator $\hat{\vec{P}}$. The eigenvalue equation

$$-i\hbar\vec{\nabla}\psi(\vec{r}) = \vec{p}\psi(\vec{r}) \tag{3.28}$$

yields the eigenfunction $\psi(\vec{r})$ corresponding to the eigenvalue $\vec{p}$ such that $|\psi(\vec{r})|^2 d^3r$ is the probability of finding the particle with a momentum $\vec{p}$ in the volume element d^3r centered about $\vec{r}$. The solution to the eigenvalue equation (3.28) is

$$\psi(\vec{r}) = Ae^{i\vec{p}\cdot\vec{r}/\hbar}, \tag{3.29}$$

where A is a normalization constant. Since $\vec{p} = \hbar\vec{k}$ is the eigenvalue of the operator $\hat{\vec{P}}$, the eigenfunction (3.29) reduces to $\psi(\vec{r}) = Ae^{i\vec{k}\cdot\vec{r}}$; hence the eigenvalue equation (3.28) becomes

$$\hat{\vec{P}}\psi(\vec{r}) = \hbar\vec{k}\psi(\vec{r}). \tag{3.30}$$

To summarize, there is a one-to-one correspondence between observables and operators (Table 3.1).

Example 3.3 (Orbital angular momentum)
Find the operator representing the classical orbital angular momentum.

Solution
The classical expression for the orbital angular momentum of a particle whose position and linear momentum are $\vec{r}$ and $\vec{p}$ is given by $\vec{L} = \vec{r}\times\vec{p} = l_x\vec{i} + l_y\vec{j} + l_z\vec{k}$, where $l_x = yp_z - zp_y$, $l_y = zp_x - xp_z$, $l_z = xp_y - yp_x$.

To find the operator representing the classical angular momentum, we need simply to replace $\vec{r}$ and $\vec{p}$ with their corresponding operators $\hat{\vec{R}}$ and $\hat{\vec{P}} = -i\hbar\vec{\nabla}$: $\hat{\vec{L}} = -i\hbar\hat{\vec{R}}\times\vec{\nabla}$. This leads to

$$\hat{L}_x = \hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y = -i\hbar\left(\hat{Y}\frac{\partial}{\partial z} - \hat{Z}\frac{\partial}{\partial y}\right), \tag{3.31}$$

$$\hat{L}_y = \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z = -i\hbar\left(\hat{Z}\frac{\partial}{\partial x} - \hat{X}\frac{\partial}{\partial z}\right), \tag{3.32}$$

$$\hat{L}_z = \hat{X}\hat{P}_y - \hat{Y}\hat{P}_x = -i\hbar\left(\hat{X}\frac{\partial}{\partial y} - \hat{Y}\frac{\partial}{\partial x}\right). \tag{3.33}$$

Recall that in classical mechanics the position and momentum components commute, $xp_x = p_x x$, and so do the components of the angular momentum, $l_x l_y = l_y l_x$. In quantum mechanics, however, this is not the case, since $\hat{X}\hat{P}_x = \hat{P}_x\hat{X} + i\hbar$ and, as will be shown in Chapter 5, $\hat{L}_x\hat{L}_y = \hat{L}_y\hat{L}_x + i\hbar\hat{L}_z$, and so on.

3.5 Measurement in Quantum Mechanics

Quantum theory is about the results of measurement; it says nothing about what might happen in the physical world outside the context of measurement. So the emphasis is on measurement.

3.5.1 How Measurements Disturb Systems

In classical physics it is possible to perform measurements on a system without disturbing it significantly. In quantum mechanics, however, the measurement process perturbs the system significantly. While carrying out measurements on classical systems, this perturbation does exist, but it is small enough that it can be neglected. In atomic and subatomic systems, however, the act of measurement induces nonnegligible or significant disturbances.

As an illustration, consider an experiment that measures the position of a hydrogenic electron. For this, we need to bombard the electron with electromagnetic radiation (photons). If we want to determine the position accurately, the wavelength of the radiation must be sufficiently short. Since the electronic orbit is of the order of 10^{-10} m, we must use a radiation whose wavelength is smaller than 10^{-10} m. That is, we need to bombard the electron with photons of energies higher than

$$h\nu = h\frac{c}{\lambda} = h\frac{3\times 10^8}{10^{-10}} \sim 10^4 \text{ eV}. \tag{3.34}$$

When such photons strike the electron, not only will they perturb it, they will knock it completely off its orbit; recall that the ionization energy of the hydrogen atom is about 13.5 eV. Thus, the mere act of measuring the position of the electron disturbs it appreciably.

Let us now discuss the general concept of measurement in quantum mechanics. *The act of measurement generally changes the state of the system.* In theory we can represent the measuring device by an operator so that, after carrying out the measurement, the system will be in one of the eigenstates of the operator. Consider a system which is in a state $|\psi\rangle$. Before measuring an observable A, the state $|\psi\rangle$ can be represented by a linear superposition of eigenstates $|\phi_n\rangle$ of the corresponding operator $\hat{A}$:

$$|\psi\rangle = \sum_n |\phi_n\rangle\langle\phi_n|\psi\rangle = \sum_n a_n|\phi_n\rangle. \tag{3.35}$$

According to Postulate 4, the act of measuring A changes the state of the system from $|\psi\rangle$ to one of the eigenstates $|\phi_n\rangle$ of the operator $\hat{A}$, and the result obtained is the eigenvalue a_n. The only *exception* to this rule is when the *system is already in one of the eigenstates of the observable being measured.* For instance, if the system is in the eigenstate $|\phi_n\rangle$, a measurement of the

observable A yields with certainty (i.e., with probability = 1) the value a_n without changing the state $|\phi_n\rangle$.

Before a measurement, we do not know in advance with certainty in which eigenstate, among the various states $|\phi_n\rangle$, a system will be after the measurement; only a probabilistic outcome is possible. Postulate 4 states that the probability of finding the system in one particular nondegenerate eigenstate $|\phi_n\rangle$ is given by

$$P_n = \frac{|\langle\phi_n|\psi\rangle|^2}{\langle\psi|\psi\rangle}. \tag{3.36}$$

Note that the wave function does not predict the results of individual measurements; it instead determines the probability distribution, $P = |\psi|^2$, over measurements on many identical systems in the same state.

Finally, we may state that quantum mechanics is the mechanics applicable to objects for which measurements necessarily interfere with the state of the system. Quantum mechanically, we cannot ignore the effects of the measuring equipment on the system, for they are important. In general, certain measurements cannot be performed without major disturbances to other properties of the quantum system. In conclusion, *it is the effects of the interference by the equipment on the system which is the essence of quantum mechanics.*

3.5.2 Expectation Values

The expectation value $\langle\hat{A}\rangle$ of $\hat{A}$ with respect to a state $|\psi\rangle$ is defined by

$$\langle\hat{A}\rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}. \tag{3.37}$$

For instance, the energy of a system is given by the expectation value of the Hamiltonian: $E = \langle\hat{H}\rangle = \langle\psi|\hat{H}|\psi\rangle/\langle\psi|\psi\rangle$.

In essence, the expectation value $\langle\hat{A}\rangle$ represents the average result of measuring $\hat{A}$ on the state $|\psi\rangle$. To see this, using the complete set of eigenvectors $|\phi_n\rangle$ of $\hat{A}$ as a basis (i.e., $\hat{A}$ is diagonal in ϕ_n), we can rewrite $\langle\hat{A}\rangle$ as follows:

$$\langle\hat{A}\rangle = \frac{1}{\langle\psi|\psi\rangle}\sum_{nm}\langle\psi|\phi_m\rangle\langle\phi_m|\hat{A}|\phi_n\rangle\langle\phi_n|\psi\rangle = \sum_n a_n\frac{|\langle\phi_n|\psi\rangle|^2}{\langle\psi|\psi\rangle}, \tag{3.38}$$

where we have used $\langle\phi_m|\hat{A}|\phi_n\rangle = a_n\delta_{nm}$. Since the quantity $|\langle\phi_n|\psi\rangle|^2/\langle\psi|\psi\rangle$ gives the probability P_n of finding the value a_n after measuring the observable A, we can indeed interpret $\langle\hat{A}\rangle$ as an *average* of a series of measurements of A:

$$\boxed{\langle\hat{A}\rangle = \sum_n a_n\frac{|\langle\phi_n|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \sum_n a_n P_n.} \tag{3.39}$$

That is, the expectation value of an observable is obtained by adding all permissible eigenvalues a_n, with each a_n multiplied by the corresponding probability P_n.

The relation (3.39), which is valid for *discrete* spectra, can be extended to a *continuous* distribution of probabilities $P(a)$ as follows:

$$\langle \hat{A} \rangle = \frac{\int_{-\infty}^{+\infty} a \left|\psi(a)\right|^2 da}{\int_{-\infty}^{+\infty} \left|\psi(a)\right|^2 da} = \int_{-\infty}^{+\infty} a\, dP(a). \tag{3.40}$$

The expectation value of an observable can be obtained physically as follows: prepare a very large number of *identical* systems each in the *same* state $|\psi\rangle$. The observable A is then measured on all these identical systems; the results of these measurements are $a_1, a_2, \ldots, a_n, \ldots$; the corresponding probabilities of occurrence are $P_1, P_2, \ldots, P_n, \ldots$. The average value of all these repeated measurements is called the expectation value of $\hat{A}$ with respect to the state $|\psi\rangle$.

Note that the process of obtaining different results when measuring the same observable on many identically prepared systems is contrary to classical physics, where these measurements must give the same outcome. In quantum mechanics, however, we can predict only the probability of obtaining a certain value for an observable.

Example 3.4

Consider a system whose state is given in terms of a complete and orthonormal set of five vectors $|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle, |\phi_4\rangle, |\phi_5\rangle$ as follows:

$$|\psi\rangle = \frac{1}{\sqrt{19}}|\phi_1\rangle + \frac{2}{\sqrt{19}}|\phi_2\rangle + \sqrt{\frac{2}{19}}|\phi_3\rangle + \sqrt{\frac{3}{19}}|\phi_4\rangle + \sqrt{\frac{5}{19}}|\phi_5\rangle,$$

where $|\phi_n\rangle$ are eigenstates to the system's Hamiltonian, $\hat{H}|\phi_n\rangle = n\varepsilon_0|\phi_n\rangle$ with $n = 1,2,3,4,5$, and where ε_0 has the dimensions of energy.

(a) If the energy is measured on a large number of identical systems that are all initially in the same state $|\psi\rangle$, what values would one obtain and with what probabilities?

(b) Find the average energy of one such system.

Solution

First, note that $|\psi\rangle$ is not normalized:

$$\langle\psi|\psi\rangle = \sum_{n=1}^{5} a_n^2 \langle\phi_n|\phi_n\rangle = \sum_{n=1}^{5} a_n^2 = \frac{1}{19} + \frac{4}{19} + \frac{2}{19} + \frac{3}{19} + \frac{5}{19} = \frac{15}{19}, \tag{3.41}$$

since $\langle\phi_j|\phi_k\rangle = \delta_{jk}$ with $j,k = 1,2,3,4,5$.

(a) Since $E_n = \langle\phi_n|\hat{H}|\phi_n\rangle = n\varepsilon_0$ $(n = 1,2,3,4,5)$, the various measurements of the energy of the system yield the values $E_1 = \varepsilon_0$, $E_2 = 2\varepsilon_0$, $E_3 = 3\varepsilon_0$, $E_4 = 4\varepsilon_0$, $E_5 = 5\varepsilon_0$ with the following probabilities:

$$P_1(E_1) = \frac{\left|\langle\phi_1|\psi\rangle\right|^2}{\langle\psi|\psi\rangle} = \left|\frac{1}{\sqrt{19}}\langle\phi_1|\phi_1\rangle\right|^2 \times \frac{19}{15} = \frac{1}{15}, \tag{3.42}$$

$$P_2(E_2) = \frac{\left|\langle\phi_2|\psi\rangle\right|^2}{\langle\psi|\psi\rangle} = \left|\frac{2}{\sqrt{19}}\langle\phi_2|\phi_2\rangle\right|^2 \times \frac{19}{15} = \frac{4}{15}, \tag{3.43}$$

$$P_3(E_3) = \frac{|\langle\phi_3|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \left|\sqrt{\frac{2}{19}}\langle\phi_3|\phi_3\rangle\right|^2 \times \frac{19}{15} = \frac{2}{15}, \tag{3.44}$$

$$P_4(E_4) = \frac{|\langle\phi_4|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \left|\sqrt{\frac{3}{19}}\langle\phi_4|\phi_4\rangle\right|^2 \times \frac{19}{15} = \frac{3}{15}, \tag{3.45}$$

and

$$P_5(E_5) = \frac{|\langle\phi_5|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \left|\sqrt{\frac{5}{19}}\langle\phi_5|\phi_5\rangle\right|^2 \times \frac{19}{15} = \frac{5}{15}, \tag{3.46}$$

(b) The average energy of a system is given by

$$E = \sum_{j=1}^{5} P_j E_j = \frac{1}{15}\varepsilon_0 + \frac{8}{15}\varepsilon_0 + \frac{6}{15}\varepsilon_0 + \frac{12}{15}\varepsilon_0 + \frac{25}{15}\varepsilon_0 = \frac{52}{15}\varepsilon_0. \tag{3.47}$$

This energy can also be obtained from the expectation value of the Hamiltonian:

$$\begin{aligned} E &= \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{19}{15}\sum_{n=1}^{5} a_n^2\langle\phi_n|\hat{H}|\phi_n\rangle = \frac{19}{15}\left(\frac{1}{19} + \frac{8}{19} + \frac{6}{19} + \frac{12}{19} + \frac{25}{19}\right)\varepsilon_0 \\ &= \frac{52}{15}\varepsilon_0, \end{aligned} \tag{3.48}$$

where the values of the coefficients a_n^2 are listed in (3.41).

3.5.3 Complete Sets of Commuting Operators (CSCO)

Two observables A and B are said to be *compatible* when their corresponding operators commute, $[\hat{A}, \hat{B}] = 0$; observables corresponding to noncommuting operators are said to be *noncompatible*.

In what follows we are going to consider the task of measuring two observables A and B on a given system. Since the act of measurement generally *perturbs* the system, the result of measuring A and B therefore depends on the *order* in which they are carried out. Measuring A first and then B leads[3] in general to results that are different from those obtained by measuring B first and then A. How does this take place?

If $\hat{A}$ and $\hat{B}$ do not commute and if the system is in an eigenstate $|\phi_n^{(a)}\rangle$ of $\hat{A}$, a measurement of A yields with certainty a value a_n, since $\hat{A}|\phi_n^{(a)}\rangle = a_n|\phi_n^{(a)}\rangle$. Then, when we measure B, the state of the system will be left in one of the eigenstates of B. If we measure A again, we will find a value which will be different from a_n. What is this new value? We cannot answer this question with certainty: only a probabilistic outcome is possible. For this, we need to expand the eigenstates of B in terms of those of A, and thus provide a probabilistic answer as to the

[3]The act of measuring A first and then B is represented by the action of product $\hat{B}\hat{A}$ of their corresponding operators on the state vector.

value of measuring A. *So if $\hat{A}$ and $\hat{B}$ do not commute, they cannot be measured simultaneously; the order in which they are measured matters.*

What happens when A and B commute? We can show that the results of their measurements will not depend on the order in which they are carried out. Before showing this, let us mention a useful theorem.

Theorem 3.1 *If two observables are compatible, their corresponding operators possess a set of common (or simultaneous) eigenstates (this theorem holds for both degenerate and nondegenerate eigenstates).*

Proof

We provide here a proof for the nondegenerate case only. If $|\phi_n\rangle$ is a nondegenerate eigenstate of $\hat{A}$, $\hat{A}|\phi_n\rangle = a_n|\phi_n\rangle$, we have

$$\langle\phi_m|[\hat{A}, \hat{B}]|\phi_n\rangle = (a_m - a_n)\langle\phi_m|\hat{B}|\phi_n\rangle = 0, \tag{3.49}$$

since $\hat{A}$ and $\hat{B}$ commute. So $\langle\phi_m|\hat{B}|\phi_n\rangle$ must vanish unless $a_n = a_m$. That is,

$$\langle\phi_m|\hat{B}|\phi_n\rangle = \langle\phi_n|\hat{B}|\phi_n\rangle \propto \delta_{nm}. \tag{3.50}$$

Hence the $|\phi_n\rangle$ are *joint* or *simultaneous* eigenstates of $\hat{A}$ and $\hat{B}$ (this completes the proof).

Denoting the simultaneous eigenstate of $\hat{A}$ and $\hat{B}$ by $|\phi^{(a)}_{n_1}, \phi^{(b)}_{n_2}\rangle$, we have

$$\hat{A}|\phi^{(a)}_{n_1}, \phi^{(b)}_{n_2}\rangle = a_{n_1}|\phi^{(a)}_{n_1}, \phi^{(b)}_{n_2}\rangle, \tag{3.51}$$

$$\hat{B}|\phi^{(a)}_{n_1}, \phi^{(b)}_{n_2}\rangle = b_{n_2}|\phi^{(a)}_{n_1}, \phi^{(b)}_{n_2}\rangle. \tag{3.52}$$

Theorem 3.1 can be generalized to the case of many mutually compatible observables A, B, C, ... These compatible observables possess a *complete* set of *joint eigenstates*

$$|\phi_n\rangle = |\phi^{(a)}_{n_1}, \phi^{(b)}_{n_2}, \phi^{(c)}_{n_3}, \ldots\rangle. \tag{3.53}$$

The completeness and orthonormality conditions of this set are

$$\sum_{n_1}\sum_{n_2}\sum_{n_3}\cdots|\phi^{(a)}_{n_1}, \phi^{(b)}_{n_2}, \phi^{(c)}_{n_3}, \ldots\rangle\langle\phi^{(a)}_{n_1}, \phi^{(b)}_{n_2}, \phi^{(c)}_{n_3}, \ldots| = 1; \tag{3.54}$$

$$\langle\phi_{n'}|\phi_n\rangle = \delta_{n'n} = \delta_{n_1'n_1}\delta_{n_2'n_2}\delta_{n_3'n_3}\cdots. \tag{3.55}$$

Let us now show why, when two observables A and B are compatible, the order in which we carry out their measurements is irrelevant. Measuring A first, we would find a value a_n and would leave the system in an eigenstate of A. According to Theorem 3.1, this eigenstate is also an eigenstate of B. Thus a measurement of B yields with certainty b_n without affecting the state of the system. In this way, if we measure A again, we obtain with certainty the same initial value a_n. Similarly, another measurement of B will yield b_n and will leave the system in the same joint eigenstate of A and B. Thus, if two observables A and B are compatible, and if the system is initially in an eigenstate of one of their operators, their measurements not only yield precise values (eigenvalues) but they will not depend on the order in which the measurements were performed. In this case, A and B are said to be *simultaneously measurable. So compatible observables can be measured simultaneously with arbitrary accuracy; noncompatible observables cannot.*

What happens if an operator, say $\hat{A}$, has *degenerate* eigenvalues? The specification of one eigenvalue does not uniquely determine the state of the system. Among the degenerate eigenstates of $\hat{A}$, only a subset of them are also eigenstates of $\hat{B}$. Thus, the set of states that are joint eigenstates of both $\hat{A}$ and $\hat{B}$ is not complete. To resolve the degeneracy, we can introduce a third operator $\hat{C}$ which commutes with both $\hat{A}$ and $\hat{B}$; then we can construct a set of joint eigenstates of $\hat{A}$, $\hat{B}$, and $\hat{C}$ that is complete. If the degeneracy persists, we may introduce a fourth operator $\hat{D}$ that commutes with the previous three and then look for their joint eigenstates which form a complete set. Continuing in this way, we will ultimately exhaust all the operators (that is, there are no more independent operators) which commute with each other. When that happens, we have then obtained a *complete set of commuting operators* (CSCO). Only then will the state of the system be specified unambiguously, for the joint eigenstates of the CSCO are determined uniquely and will form a complete set (recall that a complete set of eigenvectors of an operator is called a basis). We should, at this level, state the following definition.

Definition: A set of Hermitian operators, $\hat{A}$, $\hat{B}$, $\hat{C}$, ..., is called a CSCO if the operators mutually commute and if the set of their common eigenstates is complete and not degenerate (i.e., unique).

The complete commuting set may sometimes consist of only one operator. *Any operator with nondegenerate eigenvalues constitutes, all by itself, a CSCO.* For instance, the position operator $\hat{X}$ of a one-dimensional, spinless particle provides a complete set. Its momentum operator $\hat{P}$ is also a complete set; together, however, $\hat{X}$ and $\hat{P}$ cannot form a CSCO, for they do not commute. In three-dimensional problems, the three-coordinate position operators $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ form a CSCO; similarly, the components of the momentum operator $\hat{P}_x$, $\hat{P}_y$, and $\hat{P}_z$ also form a CSCO. In the case of spherically symmetric three-dimensional potentials, the set $\hat{H}$, $\hat{\vec{L}}^2$, $\hat{L}_z$ forms a CSCO. Note that in this case of spherical symmetry, we need three operators to form a CSCO because $\hat{H}$, $\hat{\vec{L}}^2$, and $\hat{L}_z$ are all degenerate; hence the complete and unique determination of the wave function cannot be achieved with one operator or with two.

In summary, when a given operator, say $\hat{A}$, is degenerate, the wave function cannot be determined uniquely unless we introduce one or more additional operators so as to form a complete commuting set.

3.5.4 Measurement and the Uncertainty Relations

We have seen in Chapter 2 that the uncertainty condition pertaining to the measurement of any two observables A and B is given by

$$\Delta A \Delta B \geq \frac{1}{2}|\langle[\hat{A}, \hat{B}]\rangle|, \tag{3.56}$$

where $\Delta A = \sqrt{\langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2}$.

Let us illustrate this on the joint measurement of the position and momentum observables. Since these observables are not compatible, their simultaneous measurement with infinite accuracy is not possible; that is, since $[\hat{X}, \hat{P}] = i\hbar$ there exists no state which is a simultaneous eigenstate of $\hat{X}$ and $\hat{P}$. For the case of the position and momentum operators, the relation

(3.56) yields

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \tag{3.57}$$

This condition shows that the position and momentum of a microscopic system cannot be measured with infinite accuracy both at once. If the position is measured with an uncertainty Δx, the uncertainty associated with its momentum measurement cannot be smaller than $\hbar/2\Delta x$. This is due to *the interference between the two measurements*. If we measure the position first, we perturb the system by changing its state to an eigenstate of the position operator; then the measurement of the momentum throws the system into an eigenstate of the momentum operator.

Another interesting application of the uncertainty relation (3.56) is to the orbital angular momentum of a particle. Since its components satisfy the commutator $[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z$, we obtain

$$\Delta L_x \Delta L_y \geq \frac{1}{2}\hbar|\langle \hat{L}_z \rangle|. \tag{3.58}$$

We can obtain the other two inequalities by means of a cyclic permutation of x, y, and z. If $\langle \hat{L}_z \rangle = 0$, $\hat{L}_x$ and $\hat{L}_y$ will have sharp values simultaneously. This occurs when the particle is in an s state. In fact, when a particle is in an s state, we have $\langle \hat{L}_x \rangle = \langle \hat{L}_y \rangle = \langle \hat{L}_z \rangle = 0$; hence all the components of orbital angular momentum will have sharp values simultaneously.

3.6 Time Evolution of the System's State

3.6.1 Time Evolution Operator

We want to examine here how quantum states evolve in time. That is, given the initial state $|\psi(t_0)\rangle$, how does one find the state $|\psi(t)\rangle$ at any later time t? The two states can be related by means of a linear operator $\hat{U}(t, t_0)$ such that

$$|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle \qquad (t > t_0); \tag{3.59}$$

$\hat{U}(t, t_0)$ is known as the *time evolution operator* or *propagator*. From (3.59), we infer that

$$\hat{U}(t_0, t_0) = \hat{I}, \tag{3.60}$$

where $\hat{I}$ is the unit (identity) operator.

The issue now is to find $\hat{U}(t, t_0)$. For this, we need simply to substitute (3.59) into the time-dependent Schrödinger equation (3.10):

$$i\hbar\frac{\partial}{\partial t}\left(\hat{U}(t, t_0)|\psi(t_0)\rangle\right) = \hat{H}\left(\hat{U}(t, t_0)|\psi(t_0)\rangle\right) \tag{3.61}$$

or

$$\frac{\partial \hat{U}(t, t_0)}{\partial t} = -\frac{i}{\hbar}\hat{H}\hat{U}(t, t_0). \tag{3.62}$$

The integration of this differential equation depends on whether or not the Hamiltonian depends on time. If it does not depend on time, and taking into account the initial condition (3.60), we can easily ascertain that the integration of (3.62) leads to

$$\boxed{\hat{U}(t, t_0) = e^{-i(t-t_0)\hat{H}/\hbar} \qquad \text{and} \qquad |\psi(t)\rangle = e^{-i(t-t_0)\hat{H}/\hbar}|\psi(t_0)\rangle.} \tag{3.63}$$

We will show in Section 3.7 that the operator $\hat{U}(t,t_0) = e^{-i(t-t_0)\hat{H}/\hbar}$ represents a finite time translation.

If, on the other hand, $\hat{H}$ depends on time the integration of (3.62) becomes less trivial. We will deal with this issue in Chapter 10 when we look at time-dependent potentials or at the time-dependent perturbation theory. In this chapter, and in all chapters up to Chapter 10, we will consider only Hamiltonians that do not depend on time.

Note that $\hat{U}(t,t_0)$ is a unitary operator, since

$$\hat{U}(t,t_0)\hat{U}^\dagger(t,t_0) = \hat{U}(t,t_0)\hat{U}^{-1}(t,t_0) = e^{-i(t-t_0)\hat{H}/\hbar}e^{i(t-t_0)\hat{H}/\hbar} = \hat{I} \tag{3.64}$$

or $\hat{U}^\dagger = \hat{U}^{-1}$.

3.6.2 Stationary States: Time-Independent Potentials

In the position representation, the time-dependent Schrödinger equation (3.10) for a particle of mass m moving in a time-dependent potential $\hat{V}(\vec{r},t)$ can be written as follows:

$$\boxed{i\hbar\frac{\partial\Psi(\vec{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\vec{r},t) + \hat{V}(\vec{r},t)\Psi(\vec{r},t).} \tag{3.65}$$

Now, let us consider the particular case of *time-independent* potentials: $\hat{V}(\vec{r},t) = \hat{V}(\vec{r})$. In this case the Hamiltonian operator will also be time independent, and hence the Schrödinger equation will have solutions that are *separable*, i.e., solutions that consist of a product of two functions, one depending only on $\vec{r}$ and the other only on time:

$$\Psi(\vec{r},t) = \psi(\vec{r})f(t). \tag{3.66}$$

Substituting (3.66) into (3.65) and dividing both sides by $\psi(\vec{r})f(t)$, we obtain

$$i\hbar\frac{1}{f(t)}\frac{df(t)}{dt} = \frac{1}{\psi(\vec{r})}\left[-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r}) + \hat{V}(\vec{r})\psi(\vec{r})\right]. \tag{3.67}$$

Since the left-hand side depends only on time and the right-hand side depends only on $\vec{r}$, both sides must be equal to a constant; this constant, which we denote by E, has the dimensions of energy. We can therefore break (3.67) into two separate differential equations, one depending on time only,

$$i\hbar\frac{df(t)}{dt} = Ef(t), \tag{3.68}$$

and the other on the space variable $\vec{r}$,

$$\boxed{\left[-\frac{\hbar^2}{2m}\nabla^2 + \hat{V}(\vec{r})\right]\psi(\vec{r}) = E\psi(\vec{r}).} \tag{3.69}$$

This equation is known as the *time-independent* Schrödinger equation for a particle of mass m moving in a time-independent potential $\hat{V}(\vec{r})$.

The solutions to (3.68) can be written as $f(t) = e^{-iEt/\hbar}$; hence the state (3.66) becomes

$$\Psi(\vec{r},t) = \psi(\vec{r})e^{-iEt/\hbar}. \tag{3.70}$$

This particular solution of the Schrödinger equation (3.65) for a *time-independent potential* is called a *stationary state*. Why is this state called *stationary*? The reason is obvious: the probability density is stationary, i.e., it does not depend on time:

$$\left|\Psi(\vec{r},t)\right|^2 = |\psi(\vec{r})e^{-iEt/\hbar}|^2 = \left|\psi(\vec{r})\right|^2. \tag{3.71}$$

Note that such a state has a precise value for the energy, $E = \hbar\omega$.

In summary, stationary states, which are given by the solutions of (3.69), exist only for time-independent potentials. The set of energy levels that are solutions to this equation are called the *energy spectrum* of the system. The states corresponding to discrete and continuous spectra are called *bound* and *unbound* states, respectively. We will consider these questions in detail in Chapter 4.

The most general solution to the time-dependent Schrödinger equation (3.65) can be written as an expansion in terms of the stationary states $\phi_n(\vec{r})\exp(-iE_nt/\hbar)$:

$$\Psi(\vec{r},t) = \sum_n c_n\phi_n(\vec{r})\exp\left(-\frac{iE_nt}{\hbar}\right), \tag{3.72}$$

where $c_n = \langle\phi_n|\Psi(t=0)\rangle = \int \phi_n^*(\vec{r})\psi(\vec{r})\,d^3r$. The general solution (3.72) is not a stationary state, because a linear superposition of stationary states is not necessarily a stationary state.

Remark

The time-dependent and time-independent Schrödinger equations are given in one dimension by (see (3.65) and (3.69))

$$\boxed{i\hbar\frac{\partial\Psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + \hat{V}(x,t)\Psi(x,t),} \tag{3.73}$$

$$\boxed{-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + \hat{V}(x)\psi(x) = E\psi(x).} \tag{3.74}$$

3.6.3 Schrödinger Equation and Wave Packets

Can we derive the Schrödinger equation (3.10) formally from first principles? No, we cannot; we can only postulate it. What we can do, however, is to provide an educated guess on the formal steps leading to it. *Wave packets* offer the formal tool to achieve that. We are going to show how to start from a wave packet and end up with the Schrödinger equation.

As seen in Chapter 1, the wave packet representing a particle of energy E and momentum p moving in a potential V is given by

$$\begin{aligned}\Psi(x,t) &= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty}\tilde{\phi}(p)\exp\left[\frac{i}{\hbar}(px-Et)\right]dp \\ &= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty}\tilde{\phi}(p)\exp\left[\frac{i}{\hbar}\left(px-\left(\frac{p^2}{2m}+V\right)t\right)\right]dp;\end{aligned} \tag{3.75}$$

recall that wave packets unify the corpuscular (E and p) and the wave (k and ω) features of particles: $k = p/\hbar$, $\hbar\omega = E = p^2/(2m) + V$. A partial time derivative of (3.75) yields

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty}\tilde{\phi}(p)\left(\frac{p^2}{2m}+V\right)\exp\left[\frac{i}{\hbar}\left(px-\left(\frac{p^2}{2m}+V\right)t\right)\right]dp. \tag{3.76}$$

Since $p^2/(2m) = -(\hbar^2/2m)\partial^2/\partial x^2$ and assuming that V is constant, we can take the term $-(\hbar^2/2m)\partial^2/\partial x^2 + V$ outside the integral sign, for it does not depend on p:

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V\right)\frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty}\tilde{\phi}(p)\exp\left[\frac{i}{\hbar}\left(px-\left(\frac{p^2}{2m}+V\right)t\right)\right]dp. \tag{3.77}$$

This can be written as

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V\right]\Psi(x,t). \tag{3.78}$$

Now, since this equation is valid for spatially varying potentials $V = V(x)$, we see that we have ended up with the Schrödinger equation (3.73).

3.6.4 The Conservation of Probability

Since the Hamiltonian operator is Hermitian, we can show that the norm $\langle\Psi(t)|\Psi(t)\rangle$, which is given by

$$\langle\Psi(t)|\Psi(t)\rangle = \int\left|\Psi(\vec{r},t)\right|^2 d^3r, \tag{3.79}$$

is time independent. This means, if $|\Psi(t)\rangle$ is normalized, it stays normalized for all subsequent times. This is a direct consequence of the hermiticity of $\hat{H}$.

To prove that $\langle\Psi(t)|\Psi(t)\rangle$ is constant, we need simply to show that its time derivative is zero. First, the time derivative of $\langle\Psi(t)|\Psi(t)\rangle$ is

$$\frac{d}{dt}\langle\Psi(t)|\Psi(t)\rangle = \left(\frac{d}{dt}\langle\Psi(t)|\right)|\Psi(t)\rangle + \langle\Psi(t)|\left(\frac{d|\Psi(t)\rangle}{dt}\right), \tag{3.80}$$

where $d|\Psi(t)\rangle/dt$ and $d\langle\Psi(t)|/dt$ can be obtained from (3.10):

$$\frac{d}{dt}|\Psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\Psi(t)\rangle, \tag{3.81}$$

$$\frac{d}{dt}\langle\Psi(t)| = \frac{i}{\hbar}\langle\Psi(t)|\hat{H}^\dagger = \frac{i}{\hbar}\langle\Psi(t)|\hat{H}. \tag{3.82}$$

Inserting these two equations into (3.80), we end up with

$$\frac{d}{dt}\langle\Psi(t)|\Psi(t)\rangle = \left(\frac{i}{\hbar}-\frac{i}{\hbar}\right)\langle\Psi(t)|\hat{H}|\Psi(t)\rangle = 0. \tag{3.83}$$

Thus, the probability density $\langle\Psi|\Psi\rangle$ does not evolve in time.

In what follows we are going to calculate the probability density in the position representation. For this, we need to invoke the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi(\vec{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\vec{r},t) + \hat{V}(\vec{r},t)\Psi(\vec{r},t) \tag{3.84}$$

and its complex conjugate

$$-i\hbar\frac{\partial\Psi^*(\vec{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi^*(\vec{r},t) + \hat{V}(\vec{r},t)\Psi^*(\vec{r},t). \tag{3.85}$$

Multiplying (3.84) on the left by $\Psi^*(\vec{r},t)$ and (3.85) on the right by $\Psi(\vec{r},t)$, then subtracting the two resulting equations (i.e., $\Psi^*(\vec{r},t)\times(3.84) - (3.85)\times\Psi(\vec{r},t)$), we obtain

$$\begin{aligned} i\hbar\Psi^*(\vec{r},t)\frac{\partial\Psi(\vec{r},t)}{\partial t} + i\hbar\frac{\partial\Psi^*(\vec{r},t)}{\partial t}\Psi(\vec{r},t) &= -\frac{\hbar^2}{2m}\Psi^*(\vec{r},t)\nabla^2\Psi(\vec{r},t) + \hat{V}(\vec{r},t)\Psi^*(\vec{r},t)\Psi(\vec{r},t) \\ &\quad + \frac{\hbar^2}{2m}\left(\nabla^2\Psi^*(\vec{r},t)\right)\Psi(\vec{r},t) - \hat{V}(\vec{r},t)\Psi^*(\vec{r},t)\Psi(\vec{r},t), \end{aligned} \tag{3.86}$$

which reduces to

$$i\frac{\partial}{\partial t}\left[\Psi^*(\vec{r},t)\Psi(\vec{r},t)\right] = -\frac{\hbar}{2m}\Psi^*(\vec{r},t)\nabla^2\Psi(\vec{r},t) + \frac{\hbar}{2m}\left(\nabla^2\Psi^*(\vec{r},t)\right)\Psi(\vec{r},t). \tag{3.87}$$

Using the identity

$$-\Psi^*(\vec{r},t)\nabla^2\Psi(\vec{r},t) + \left(\nabla^2\Psi^*(\vec{r},t)\right)\Psi(\vec{r},t) = \vec{\nabla}\cdot\left(-\Psi^*\vec{\nabla}\Psi + \Psi\vec{\nabla}\Psi^*\right), \tag{3.88}$$

we can reduce (3.87) to

$$\frac{\partial}{\partial t}\left[\Psi^*(\vec{r},t)\Psi(\vec{r},t)\right] + \frac{i\hbar}{2m}\vec{\nabla}\cdot\left(\Psi\vec{\nabla}\Psi^* - \Psi^*\vec{\nabla}\Psi\right) = 0, \tag{3.89}$$

or to

$$\boxed{\frac{\partial\rho(\vec{r},t)}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0,} \tag{3.90}$$

where $\rho(\vec{r},t)$ and $\vec{J}$ are given by

$$\boxed{\rho(\vec{r},t) = \Psi^*(\vec{r},t)\Psi(\vec{r},t), \qquad \vec{J}(\vec{r},t) = \frac{i\hbar}{2m}\left(\Psi\vec{\nabla}\Psi^* - \Psi^*\vec{\nabla}\Psi\right);} \tag{3.91}$$

$\rho(\vec{r},t)$ is called the *probability density*, while $\vec{J}(\vec{r},t)$ is the *probability current density*, or simply the *current density*, or even the particle density flux. By analogy with charge conservation in electrodynamics, equation (3.90) is interpreted as the *conservation of probability*.

Let us find the relationship between the density operators $\hat{\rho}(t)$ and $\hat{\rho}(t_0)$. Since $|\Psi(t)\rangle = \hat{U}(t,t_0)|\Psi(t_0)\rangle$ and $\langle\Psi(t)| = \langle\Psi(t_0)|\hat{U}^\dagger(t,t_0)$, we have

$$\hat{\rho}(t) = |\Psi(t)\rangle\langle\Psi(t)| = \hat{U}(t,t_0)|\Psi(0)\rangle\langle\Psi(0)|\hat{U}^\dagger(t,t_0). \tag{3.92}$$

This is known as the *density operator* for the state $|\Psi(t)\rangle$. Hence knowing $\hat{\rho}(t_0)$ we can calculate $\hat{\rho}(t)$ as follows:

$$\hat{\rho}(t) = \hat{U}(t,t_0)\hat{\rho}(t_0)\hat{U}^\dagger(t,t_0). \tag{3.93}$$

3.6.5 Time Evolution of Expectation Values

We want to look here at the time dependence of the expectation value of a linear operator; if the state $|\Psi(t)\rangle$ is normalized, the expectation value is given by

$$\langle \hat{A} \rangle = \langle \Psi(t)|\hat{A}|\Psi(t)\rangle. \tag{3.94}$$

Using (3.81) and (3.82), we can write $d\langle \hat{A}\rangle/dt$ as follows:

$$\frac{d}{dt}\langle \hat{A} \rangle = \frac{1}{i\hbar}\langle \Psi(t)|\hat{A}\hat{H} - \hat{H}\hat{A}|\Psi(t)\rangle + \langle \Psi(t)|\frac{\partial A}{\partial t}|\Psi(t)\rangle \tag{3.95}$$

or

$$\boxed{\frac{d}{dt}\langle \hat{A} \rangle = \frac{1}{i\hbar}\langle [\hat{A}, \hat{H}]\rangle + \langle \frac{\partial \hat{A}}{\partial t}\rangle.} \tag{3.96}$$

Two important results stem from this relation. First, if the observable A does not depend explicitly on time, the term $\partial\hat{A}/\partial t$ will vanish, so the rate of change of the expectation value of $\hat{A}$ is given by $\langle [\hat{A}, \hat{H}]\rangle/i\hbar$. Second, besides not depending explicitly on time, if the observable A commutes with the Hamiltonian, the quantity $d\langle \hat{A}\rangle/dt$ will then be zero; hence the expectation value $\langle \hat{A}\rangle$ will be constant in time. So if $\hat{A}$ commutes with the Hamiltonian and is not dependent on time, the observable A is said to be a *constant of the motion*; that is, *the expectation value of an operator that does not depend on time and that commutes with the Hamiltonian is constant in time*:

$$\boxed{\text{If} \quad [\hat{H}, \hat{A}] = 0 \quad \text{and} \quad \frac{\partial \hat{A}}{\partial t} = 0 \implies \frac{d\langle \hat{A}\rangle}{dt} = 0 \implies \langle \hat{A}\rangle = \text{constant}.} \tag{3.97}$$

For instance, we can verify that the energy, the linear momentum, and the angular momentum of an isolated system are conserved: $d\langle \hat{H}\rangle/dt = 0$, $d\langle \hat{\vec{P}}\rangle/dt = 0$, and $d\langle \hat{\vec{L}}\rangle/dt = 0$. This implies that the expectation values of $\hat{H}$, $\hat{\vec{P}}$, and $\hat{\vec{L}}$ are constant. Recall from classical physics that the conservation of energy, linear momentum, and angular momentum are consequences of the following symmetries, respectively: homogeneity of time, homogeneity of space, and isotropy of space. We will show in the following section that these symmetries are associated, respectively, with invariances in time translation, space translation, and space rotation.

As an example, let us consider the time evolution of the expectation value of the density operator $\hat{\rho}(t) = |\Psi(t)\rangle\langle\Psi(t)|$; see (3.92). From (3.10), which leads to $\partial|\Psi(t)\rangle/\partial t = (1/i\hbar)\hat{H}|\Psi(t)\rangle$ and $\partial\langle\Psi(t)|/\partial t = -(1/i\hbar)\langle\Psi(t)|\hat{H}$, we have

$$\frac{\partial \hat{\rho}(t)}{\partial t} = \frac{1}{i\hbar}\hat{H}|\Psi(t)\rangle\langle\Psi(t)| - \frac{1}{i\hbar}|\Psi(t)\rangle\langle\Psi(t)|\hat{H} = -\frac{1}{i\hbar}[\hat{\rho}(t), \hat{H}]. \tag{3.98}$$

A substitution of this relation into (3.96) leads to

$$\frac{d}{dt}\langle \hat{\rho}(t)\rangle = \frac{1}{i\hbar}\langle [\hat{\rho}(t), \hat{H}]\rangle + \langle \frac{\partial \hat{\rho}(t)}{\partial t}\rangle = \frac{1}{i\hbar}\langle [\hat{\rho}(t), \hat{H}]\rangle - \frac{1}{i\hbar}\langle [\hat{\rho}(t), \hat{H}]\rangle = 0. \tag{3.99}$$

So the density operator is a constant of the motion. In fact, we can easily show that

$$\begin{aligned} \langle [\hat{\rho}(t), \hat{H}]\rangle &= \langle \Psi(t)|[|\Psi(t)\rangle\langle\Psi(t)|, \hat{H}]|\Psi(t)\rangle \\ &= \langle \Psi(t)|\Psi(t)\rangle\langle\Psi(t)|\hat{H}|\Psi(t)\rangle - \langle \Psi(t)|\hat{H}|\Psi(t)\rangle\langle\Psi(t)|\Psi(t)\rangle \\ &= 0, \end{aligned} \tag{3.100}$$

which, when combined with (3.98), yields $\langle \partial \hat{p}(t)/\partial t \rangle = 0$.

Finally, we should note that the constants of motion are nothing but observables that can be measured simultaneously with the energy to arbitrary accuracy. If a system has a complete set of commuting operators (CSCO), the number of these operators is given by the total number of constants of the motion.

3.7 Symmetries and Conservation Laws

We are interested here in symmetries that leave the Hamiltonian of an *isolated* system invariant. We will show that for each such symmetry there corresponds an observable which is a constant of the motion. The invariance principles relevant to our study are the time translation invariance and the space translation invariance. We may recall from classical physics that whenever a system is invariant under space translations, its total momentum is conserved; and whenever it is invariant under rotations, its total angular momentum is also conserved.

To prepare the stage for symmetries and conservation laws in quantum mechanics, we are going to examine the properties of infinitesimal and finite unitary transformations that are most essential to these invariance principles.

3.7.1 Infinitesimal Unitary Transformations

In Chapter 2 we saw that the transformations of a state vector $|\psi\rangle$ and an operator $\hat{A}$ under an infinitesimal unitary transformation $U_\varepsilon(\hat{G}) = \hat{I} + i\varepsilon\hat{G}$ are given by

$$|\psi'\rangle = (\hat{I} + i\varepsilon\hat{G})|\psi\rangle = |\psi\rangle + \delta|\psi\rangle, \tag{3.101}$$

$$\hat{A}' = (\hat{I} + i\varepsilon\hat{G})\hat{A}(\hat{I} - i\varepsilon\hat{G}) \simeq \hat{A} + i\varepsilon[\hat{G}, \hat{A}], \tag{3.102}$$

where ε and $\hat{G}$ are called the parameter and the generator of the transformation, respectively.

Let us consider two important applications of infinitesimal unitary transformations: time and space translations.

3.7.1.1 Time Translations: $\hat{G} = \hat{H}/\hbar$

The application of $\hat{U}_{\delta t}(\hat{H}) = \hat{I} + (i/\hbar)\delta t\,\hat{H}$ on a state $|\psi(t)\rangle$ gives

$$\left(\hat{I} + \frac{i}{\hbar}\,\delta t\,\hat{H}\right)|\psi(t)\rangle = |\psi(t)\rangle + \left(\frac{i}{\hbar}\delta t\right)\hat{H}|\psi(t)\rangle. \tag{3.103}$$

Since $\hat{H}|\psi(t)\rangle = i\hbar\partial|\psi(t)\rangle/\partial t$ we have

$$\boxed{\left(\hat{I} + \frac{i}{\hbar}\,\delta t\,\hat{H}\right)|\psi(t)\rangle = |\psi(t)\rangle - \delta t\frac{\partial|\psi(t)\rangle}{\partial t} \simeq |\psi(t - \delta t)\rangle,} \tag{3.104}$$

because $|\psi(t)\rangle - \delta t\,\partial|\psi(t)\rangle/\partial t$ is nothing but the first-order Taylor expansion of $|\psi(t - \delta t)\rangle$. We conclude from (3.104) that the application of $\hat{U}_{\delta t}(\hat{H})$ to $|\psi(t)\rangle$ generates a state $|\psi(t - \delta t)\rangle$ which consists simply of a *time translation* of $|\psi(t)\rangle$ by an amount equal to δt. *The Hamiltonian in* $(\hat{I} + (i/\hbar)\delta t\,\hat{H})$ *is thus the generator of infinitesimal time translations.* Note that this translation preserves the shape of the state $|\psi(t)\rangle$, for its overall shape is merely translated in time by δt.

3.7.1.2 Spatial Translations: $\hat{G} = \hat{P}_x/\hbar$

The application of $\hat{U}_\varepsilon(\hat{P}_x) = \hat{I} + (i/\hbar)\varepsilon\hat{P}_x$ to $\psi(x)$ gives

$$\left(\hat{I} + \frac{i}{\hbar}\varepsilon\hat{P}_x\right)\psi(x) = \psi(x) + \left(\frac{i}{\hbar}\varepsilon\right)\hat{P}_x\psi(x). \tag{3.105}$$

Since $\hat{P}_x = -i\hbar\partial/\partial x$ and since the first-order Taylor expansion of $\psi(x+\varepsilon)$ is given by $\psi(x+\varepsilon) = \psi(x) + \varepsilon\partial\psi(x)/\partial x$, we have

$$\boxed{\left(\hat{I} + \frac{i}{\hbar}\varepsilon\hat{P}_x\right)\psi(x) = \psi(x) + \varepsilon\frac{\partial\psi(x)}{\partial x} \simeq \psi(x+\varepsilon).} \tag{3.106}$$

So, when $\hat{U}_\varepsilon(\hat{P}_x)$ acts on a wave function, it translates it spatially by an amount equal to ε.

Using $[\hat{X}, \hat{P}_x] = i\hbar$ we infer from (3.102) that the position operator $\hat{X}$ transforms as follows:

$$\hat{X}' = \left(\hat{I} + \frac{i}{\hbar}\varepsilon\hat{P}_x\right)\hat{X}\left(\hat{I} - \frac{i}{\hbar}\varepsilon\hat{P}_x\right) \simeq \hat{X} + \frac{i}{\hbar}\varepsilon[\hat{P}_x, \hat{X}] = \hat{X} + \varepsilon. \tag{3.107}$$

The relations (3.106) and (3.107) show that the *linear momentum operator in* $(\hat{I} + (i/\hbar)\varepsilon\hat{P}_x)$ *is a generator of infinitesimal spatial translations.*

3.7.2 Finite Unitary Transformations

In Chapter 2 we saw that a *finite* unitary transformation can be constructed by performing a succession of infinitesimal transformations. For instance, by applying a single infinitesimal time translation N times in steps of τ/N, we can generate a finite time translation

$$\hat{U}_\tau(\hat{H}) = \lim_{N\to+\infty}\prod_{k=1}^{N}\left(\hat{I} + \frac{i}{\hbar}\frac{\tau}{N}\hat{H}\right) = \lim_{N\to+\infty}\left(\hat{I} + \frac{i}{\hbar}\tau\hat{H}\right)^N = \exp\left(\frac{i}{\hbar}\tau\hat{H}\right), \tag{3.108}$$

where the Hamiltonian is the generator of finite time translations. We should note that the time evolution operator $\hat{U}(t, t_0) = e^{-i(t-t_0)\hat{H}/\hbar}$, displayed in (3.63), represents a finite unitary transformation where $\hat{H}$ is the generator of the time translation.

By analogy with (3.104) we can show that the application of $\hat{U}_\tau(\hat{H})$ to $|\psi(t)\rangle$ yields

$$\boxed{\hat{U}_\tau(\hat{H})|\psi(t)\rangle = \exp\left(\frac{i}{\hbar}\tau\hat{H}\right)|\psi(t)\rangle = |\psi(t-\tau)\rangle,} \tag{3.109}$$

where $|\psi(t-\tau)\rangle$ is merely a time translation of $|\psi(t)\rangle$.

Similarly, we can infer from (3.106) that the application of $\hat{U}_{\vec{a}}(\hat{\vec{P}}) = \exp(i\vec{a}\cdot\hat{\vec{P}}/\hbar)$ to a wave function causes it to be translated in space by a vector $\vec{a}$:

$$\boxed{\hat{U}_{\vec{a}}(\hat{\vec{P}})\psi(\vec{r}) = \exp\left(\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right)\psi(\vec{r}) = \psi(\vec{r}+\vec{a}).} \tag{3.110}$$

To calculate the transformed position vector operator $\hat{\vec{R}}'$, let us invoke a relation we derived in Chapter 2:

$$\hat{A}' = e^{i\alpha\hat{G}}\hat{A}e^{-i\alpha\hat{G}} = \hat{A} + i\alpha[\hat{G},\hat{A}] + \frac{(i\alpha)^2}{2!}[\hat{G},[\hat{G},\hat{A}]] + \frac{(i\alpha)^3}{3!}[\hat{G},[\hat{G},[\hat{G},\hat{A}]]] + \cdots . \tag{3.111}$$

An application of this relation to the spatial translation operator $\hat{U}_a(\hat{\vec{P}})$ yields

$$\hat{\vec{R}}' = \exp\left(\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right)\hat{\vec{R}}\exp\left(-\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right) = \hat{\vec{R}} + \frac{i}{\hbar}[\vec{a}\cdot\hat{\vec{P}},\hat{\vec{R}}] = \hat{\vec{R}} + \vec{a}. \tag{3.112}$$

In deriving this, we have used the fact that $[\vec{a}\cdot\hat{\vec{P}},\hat{\vec{R}}] = -i\hbar\vec{a}$ and that the other commutators are zero, notably $[\vec{a}\cdot\hat{\vec{P}},[\vec{a}\cdot\hat{\vec{P}},\hat{\vec{R}}]] = 0$. From (3.110) and (3.112), we see that *the linear momentum in* $\exp(i\vec{a}\cdot\hat{\vec{P}}/\hbar)$ *is a generator of finite spatial translations.*

3.7.3 Symmetries and Conservation Laws

We want to show here that every invariance principle of $\hat{H}$ is connected with a conservation law.

The Hamiltonian of a system transforms under a unitary transformation $e^{i\alpha\hat{G}}$ as follows; see (3.111):

$$\hat{H}' = e^{i\alpha\hat{G}}\hat{H}e^{-i\alpha\hat{G}} = \hat{H} + i\alpha[\hat{G},\hat{H}] + \frac{(i\alpha)^2}{2!}[\hat{G},[\hat{G},\hat{H}]] + \frac{(i\alpha)^3}{3!}[\hat{G},[\hat{G},[\hat{G},\hat{H}]]] + \cdots . \tag{3.113}$$

If $\hat{H}$ commutes with $\hat{G}$, it also commutes with the unitary transformation $\hat{U}_a(\hat{G}) = e^{i\alpha\hat{G}}$. In this case we may infer two important conclusions. On the one hand, there is an *invariance principle*: the Hamiltonian is invariant under the transformation $\hat{U}_a(\hat{G})$, since

$$\hat{H}' = e^{i\alpha\hat{G}}\hat{H}e^{-i\alpha\hat{G}} = e^{i\alpha\hat{G}}e^{-i\alpha\hat{G}}\hat{H} = \hat{H}. \tag{3.114}$$

On the other hand, if in addition to $[\hat{G},\hat{H}] = 0$, the operator $\hat{G}$ does not depend on time explicitly, there is a *conservation law*: equation (3.96) shows that $\hat{G}$ is a *constant of the motion*, since

$$\frac{d}{dt}\langle\hat{G}\rangle = \frac{1}{i\hbar}\langle[\hat{G},\hat{H}]\rangle + \langle\frac{\partial\hat{G}}{\partial t}\rangle = 0. \tag{3.115}$$

We say that $\hat{G}$ is conserved.

So whenever the Hamiltonian is invariant under a unitary transformation, the generator of the transformation is conserved. We may say, in general, that *for every invariance symmetry of the Hamiltonian, there corresponds a conservation law.*

3.7.3.1 Conservation of Energy and Linear Momentum

Let us consider two interesting applications pertaining to the invariance of the Hamiltonian of an *isolated* system with respect to time translations and to space translations. First, let us consider time translations. As shown in (3.63), time translations are generated in the case

of time-independent Hamiltonians by the evolution operator $\hat{U}(t,t_0) = e^{-i(t-t_0)\hat{H}/\hbar}$. Since $\hat{H}$ commutes with the generator of the time translation (which is given by $\hat{H}$ itself), it is invariant under time translations. *As $\hat{H}$ is invariant under time translations, the energy of an isolated system is conserved.* We should note that if the system is invariant under time translations, this means there is a symmetry of time homogeneity. Time homogeneity implies that the time-displaced state $\psi(t-\tau)$, like $\psi(t)$, satisfies the Schrödinger equation.

The second application pertains to the spatial translations, or to transformations under $\hat{U}_{\vec{a}}(\hat{\vec{P}}) = \exp(i\vec{a}\cdot\hat{\vec{P}}/\hbar)$, of an isolated system. The linear momentum is invariant under $\hat{U}_{\vec{a}}(\hat{\vec{P}})$ and the position operator transforms according to (3.112):

$$\hat{\vec{P}}' = \hat{\vec{P}}, \qquad \hat{\vec{R}}' = \hat{\vec{R}} + \vec{a}. \tag{3.116}$$

For instance, since the Hamiltonian of a free particle does not depend on the coordinates, it commutes with the linear momentum $[\hat{H},\hat{\vec{P}}] = 0$. The Hamiltonian is then invariant under spatial translations, since

$$\hat{H}' = \exp\left(\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right)\hat{H}\exp\left(-\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right) = \exp\left(\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right)\exp\left(-\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right)\hat{H} = \hat{H}. \tag{3.117}$$

Since $[\hat{H},\hat{\vec{P}}] = 0$ and since the linear momentum operator does not depend explicitly on time, we infer from (3.96) that $\hat{\vec{P}}$ is a constant of the motion, since

$$\frac{d}{dt}\langle\hat{\vec{P}}\rangle = \frac{1}{i\hbar}\langle[\hat{\vec{P}},\hat{H}]\rangle + \langle\frac{\partial\hat{\vec{P}}}{\partial t}\rangle = 0. \tag{3.118}$$

So if $[\hat{H},\hat{\vec{P}}] = 0$ the *Hamiltonian will be invariant under spatial translations and the linear momentum will be conserved.* A more general case where the linear momentum is a constant of the motion is provided by an isolated system, for its total linear momentum is conserved. Note that the invariance of the system under spatial translations means there is a symmetry of spatial homogeneity. The requirement for the homogeneity of space implies that the spatially displaced wave function $\psi(\vec{r}+\vec{a})$, much like $\psi(\vec{r})$, satisfies the Schrödinger equation.

In summary, the symmetry of time homogeneity gives rise to the conservation of energy, whereas the symmetry of space homogeneity gives rise to the conservation of linear momentum.

In Chapter 7 we will see that the symmetry of space isotropy, or the invariance of the Hamiltonian with respect to space rotations, leads to conservation of the angular momentum.

Parity operator

The unitary transformations we have considered so far, time translations and space translations, are *continuous*. We may consider now a *discrete* unitary transformation, the *parity*. As seen in Chapter 2, the parity transformation consists of an inversion or reflection through the origin of the coordinate system:

$$\hat{P}\psi(\vec{r}) = \psi(-\vec{r}). \tag{3.119}$$

If the parity operator commutes with the system's Hamiltonian,

$$[\hat{H},\hat{P}] = 0, \tag{3.120}$$

the *parity will be conserved*, and hence a constant of the motion. In this case the Hamiltonian and the parity operator have simultaneous eigenstates. For instance, we will see in Chapter 4 that the wave functions of a particle moving in a symmetric potential, $\hat{V}(\vec{r}) = \hat{V}(-\vec{r})$, have definite parities: they can be only even or odd. Similarly, we can ascertain that the parity of an isolated system is a constant of the motion.

3.8 Connecting Quantum to Classical Mechanics

3.8.1 Poisson Brackets and Commutators

To establish a connection between quantum mechanics and classical mechanics, we may look at the time evolution of *dynamical observable*.

Before describing the time evolution of a dynamical variable within the context of classical mechanics, let us review the main ideas of the mathematical tool relevant to this description, the *Poisson bracket*. The *Poisson bracket* between two *dynamical observables* $A(q, p)$ and $B(q, p)$, where $q = (q_1, q_2, \cdots, q_n)$ and $p = (p_1, p_2, \cdots, p_n)$, is defined in terms of the *generalized coordinates*[4] q_j and the momenta p_j of the system as follows:

$$\{A, B\} = \sum_{j=1}^{n} \left(\frac{\partial A}{\partial q_j} \frac{\partial B}{\partial p_j} - \frac{\partial A}{\partial p_j} \frac{\partial B}{\partial q_j} \right). \tag{3.121}$$

Since the variables q_j are independent of p_j, we have $\partial q_j/\partial p_k = 0$, $\partial p_j/\partial q_k = 0$; thus we can show that

$$\{q_j, q_k\} = \{p_j, p_k\} = 0, \qquad \{q_j, p_k\} = \delta_{jk}. \tag{3.122}$$

Using (3.121) we can easily infer the following properties of the Poisson brackets:

- Antisymmetry
$$\{A, B\} = -\{B, A\} \tag{3.123}$$
- Linearity
$$\{A, \alpha B + \beta C + \gamma D + \cdots\} = \alpha\{A, B\} + \beta\{A, C\} + \gamma\{A, D\} + \cdots \tag{3.124}$$
- Complex conjugate
$$\{A, B\}^* = \{A^*, B^*\} \tag{3.125}$$
- Distributivity
$$\{A, BC\} = \{A, B\}C + B\{A, C\}, \qquad \{AB, C\} = A\{B, C\} + \{A, C\}B \tag{3.126}$$
- Jacobi identity
$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0 \tag{3.127}$$
- Using $df^n(x)/dx = nf^{n-1}(x)df(x)/dx$, we can show that
$$\{A, B^n\} = nB^{n-1}\{A, B\}, \qquad \{A^n, B\} = nA^{n-1}\{A, B\} \tag{3.128}$$

[4]For a short presentation about the generalized coordinates, see Chapter 12 (Page 980).

These properties are similar to the properties of the quantum mechanical commutators seen in Chapter 2.

A dynamical observable A is simply any function of $q = (q_1, q_2, \cdots, q_n)$, $p = (p_1, p_2, \cdots, p_n)$, and possibly time, t; i.e., $A = A(q, p, t)$. The total time derivative of a A is given by

$$\frac{dA}{dt} = \sum_{j=1}^{n}\left(\frac{\partial A}{\partial q_j}\frac{\partial q_j}{\partial t} + \frac{\partial A}{\partial p_j}\frac{\partial p_j}{\partial t}\right) + \frac{\partial A}{\partial t} = \sum_{j=1}^{n}\left(\frac{\partial A}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial A}{\partial p_j}\frac{\partial H}{\partial q_j}\right) + \frac{\partial A}{\partial t}; \tag{3.129}$$

in deriving this relation we have used the *Hamilton equations* of classical mechanics:

$$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}, \qquad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}, \tag{3.130}$$

where H is the Hamiltonian of the system. The total time evolution of a dynamical variable A is thus given by the following equation of motion:

$$\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t}. \tag{3.131}$$

If the observable A depends on t, it is said to have an *explicit time-dependence*; i.e., $A = A(q, p, t)$. However, if A does not depend explicitly on time, but depends on t only through $q_j(t)$ and $p_j(t)$, the observable is said to have an *implicit time-dependence*; i.e., $A = A(q, p)$. In short, whereas the first terms $\{A, H\}$ in (3.131) represents the implicit time-dependence of A, the last term $\partial A/\partial t$ represents the explicit time-dependence. As such, Equation (3.131) shows that, if A does not depend explicitly on time, we have $\partial A/\partial t = 0$ and, hence, dA/dt is given by

$$\text{If} \quad \frac{\partial A}{\partial t} = 0 \quad \Longrightarrow \quad \frac{dA}{dt} = \{A, H\}. \tag{3.132}$$

Now, if $dA/dt = 0$ or $\{A, H\} = 0$, the dynamical observable A is said to be a *constant of the motion*. That is, an observable A is a constant of the motion if its Poisson bracket with the Hamiltonian is zero.

Comparing the classical relation (3.131) with its quantum mechanical counterpart (3.96),

$$\frac{d}{dt}\langle\hat{A}\rangle = \frac{1}{i\hbar}\langle[\hat{A}, \hat{H}]\rangle + \langle\frac{\partial\hat{A}}{\partial t}\rangle, \tag{3.133}$$

we see that they are identical only if we identify the Poisson bracket $\{A, H\}$ with the commutator $[\hat{A}, \hat{H}]/(i\hbar)$. We may thus infer the following general rule. The Poisson bracket of any pair of classical variables can be obtained from the commutator between the corresponding pair of quantum operators by dividing it by $i\hbar$:

$$\boxed{\frac{1}{i\hbar}[\hat{A}, \hat{B}] \longrightarrow \{A, B\}_{classical}.} \tag{3.134}$$

Note that the expressions of classical mechanics can be derived from their quantum counterparts, but the opposite is not possible. That is, dividing quantum mechanical expressions by $i\hbar$ leads to their classical analog, but multiplying classical mechanical expressions by $i\hbar$ doesn't necessarily lead to their quantum counterparts.

Notice that, if we set $A = H$ in Equation (3.131), we have $\{A, H\} = \{H, H\} \equiv 0$, and hence

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}. \tag{3.135}$$

So, if the Hamiltonian does not depend explicitly on time, we have $\partial H/\partial t = 0$ and, hence, the Hamiltonian is conserved, $dH/dt = 0$. This is the general law of *energy conservation*. Evidently, the energy is not conserved if the Hamiltonian depends explicitly on time; this situation occurs when the system interacts with external, time-dependent potentials (or forces).

Example 3.5

(a) Evaluate the Poisson bracket $\{x, p\}$ between the position, x, and momentum, p, variables.

(b) Compare the commutator $\left[\hat{X}, \hat{P}\right]$ with Poisson bracket $\{x, p\}$ calculated in Part (a).

Solution

(a) Applying the general relation

$$\{A, B\} = \sum_j \left(\frac{\partial A}{\partial x_j} \frac{\partial B}{\partial p_j} - \frac{\partial A}{\partial p_j} \frac{\partial B}{\partial x_j} \right) \tag{3.136}$$

to x and p, we can readily evaluate the given Poisson bracket:

$$\begin{aligned} \{x, p\} &= \frac{\partial(x)}{\partial x}\frac{\partial(p)}{\partial p} - \frac{\partial(x)}{\partial p}\frac{\partial(p)}{\partial x} \\ &= \frac{\partial(x)}{\partial x}\frac{\partial(p)}{\partial p} \\ &= 1. \end{aligned} \tag{3.137}$$

(b) Using the fact that $[\hat{X}, \hat{P}] = i\hbar$, we see that

$$\frac{1}{i\hbar}[\hat{X}, \hat{P}] = 1, \tag{3.138}$$

which is equal to the Poisson bracket (3.137); that is,

$$\frac{1}{i\hbar}[\hat{X}, \hat{P}] = \{x, p\}_{classical} = 1. \tag{3.139}$$

This result is in agreement with Eq. (3.134).

3.8.2 The Ehrenfest Theorem

If quantum mechanics is to be more general than classical mechanics, it must contain classical mechanics as a limiting case. To illustrate this idea, let us look at the time evolution of the

expectation values of the position and momentum operators, $\hat{\vec{R}}$ and $\hat{\vec{P}}$, of a particle moving in a potential $\hat{V}(\vec{r})$, and then compare these relations with their classical counterparts.

Since the position and the momentum observables do not depend explicitly on time, within the context of wave mechanics, the terms $\langle \partial\hat{\vec{R}}/\partial t\rangle$ and $\langle \partial\hat{\vec{P}}/\partial t\rangle$ are zero. Hence, inserting $\hat{H} = \hat{\vec{P}}^2/(2m) + \hat{V}(\hat{\vec{R}}, t)$ into (3.96) and using the fact that $\hat{\vec{R}}$ commutes with $\hat{V}(\hat{\vec{R}}, t)$, we can write

$$\frac{d}{dt}\langle\hat{\vec{R}}\rangle = \frac{1}{i\hbar}\langle[\hat{\vec{R}}, \hat{H}]\rangle = \frac{1}{i\hbar}\langle[\hat{\vec{R}}, \frac{\hat{\vec{P}}^2}{2m} + \hat{V}(\hat{\vec{R}}, t)]\rangle = \frac{1}{2im\hbar}\langle[\hat{\vec{R}}, \hat{\vec{P}}^2]\rangle. \tag{3.140}$$

Since

$$[\hat{\vec{R}}, \hat{\vec{P}}^2] = 2i\hbar\hat{\vec{P}}, \tag{3.141}$$

we have

$$\boxed{\frac{d}{dt}\langle\hat{\vec{R}}\rangle = \frac{1}{m}\langle\hat{\vec{P}}\rangle.} \tag{3.142}$$

As for $d\langle\hat{\vec{P}}\rangle/dt$, we can infer its expression from a treatment analogous to $d\langle\hat{\vec{R}}\rangle/dt$. Using

$$[\hat{\vec{P}}, \hat{V}(\hat{\vec{R}}, t)] = -i\hbar\vec{\nabla}\hat{V}(\vec{R}, t), \tag{3.143}$$

we can write

$$\boxed{\frac{d}{dt}\langle\hat{\vec{P}}\rangle = \frac{1}{i\hbar}\langle[\hat{\vec{P}}, \hat{V}(\hat{\vec{R}}, t)]\rangle = -\langle\vec{\nabla}\hat{V}(\hat{\vec{R}}, t)\rangle.} \tag{3.144}$$

The two relations (3.142) and (3.144), expressing the time evolution of the expectation values of the position and momentum operators, are known as the *Ehrenfest theorem*, or Ehrenfest equations. Their respective forms are reminiscent of the Hamilton–Jacobi equations of classical mechanics,

$$\frac{d\vec{r}}{dt} = \frac{\vec{p}}{m}, \qquad \frac{d\vec{p}}{dt} = -\vec{\nabla}V(\vec{r}), \tag{3.145}$$

which reduce to Newton's equation of motion for a *classical* particle of mass m, position $\vec{r}$, and momentum $\vec{p}$:

$$\frac{d\vec{p}}{dt} = m\frac{d^2\vec{r}}{dt^2} = -\vec{\nabla}V(\vec{r}). \tag{3.146}$$

Notice $\hbar$ has completely disappeared in the Ehrenfest equations (3.142) and (3.144). These two equations certainly establish a connection between quantum mechanics and classical mechanics. We can, within this context, view the center of the wave packet as moving like a classical particle when subject to a potential $V(\vec{r})$.

3.8.3 Quantum Mechanics and Classical Mechanics

In Chapter 1 we focused mainly on those experimental observations which confirm the failure of classical physics at the microscopic level. We should bear in mind, however, that classical physics works perfectly well within the realm of the macroscopic world. Thus, if the theory of quantum mechanics is to be considered more general than classical physics, it must yield accurate results not only on the microscopic scale but at the classical limit as well.

How does one decide on when to use classical or quantum mechanics to describe the motion of a given system? That is, how do we know when a classical description is good enough or when a quantum description becomes a must? The answer is provided by comparing the size of those quantities of the system that have the dimensions of an action with the Planck constant, h. Since, as shown in (3.134), the quantum relations are characterized by h, we can state that if the value of the action of a system is too large compared to h, this system can be accurately described by means of classical physics. Otherwise, the use of a quantal description becomes unavoidable. One should recall that, for microscopic systems, the size of action variables is of the order of h; for instance, the angular momentum of the hydrogen atom is $L = n\hbar$, where n is finite.

Another equivalent way of defining the classical limit is by means of "*length*." Since $\lambda = h/p$ the classical domain can be specified by the limit $\lambda \to 0$. This means that, when the de Broglie wavelength of a system is too small compared to its size, the system can be described accurately by means of classical physics.

In summary, the classical limit can be described as the limit $h \to 0$ or, equivalently, as the limit $\lambda \to 0$. In these limits the results of quantum mechanics should be similar to those of classical physics:

$$\lim_{h \to 0} \text{Quantum Mechanics} \longrightarrow \text{Classical Mechanics}, \tag{3.147}$$

$$\lim_{\lambda \to 0} \text{Quantum Mechanics} \longrightarrow \text{Classical Mechanics}. \tag{3.148}$$

Classical mechanics can thus be regarded as the short wavelength limit of quantum mechanics. In this way, quantum mechanics contains classical mechanics as a limiting case. So, in the limit of $h \to 0$ or $\lambda \to 0$, quantum dynamical quantities should have, as proposed by Bohr, a one-to-one correspondence with their classical counterparts. This is the essence of the *correspondence principle.*

But how does one reconcile, in the classical limit, the probabilistic nature of quantum mechanics with the determinism of classical physics? The answer is quite straightforward: quantum *fluctuations* must become negligible or even vanish when $h \to 0$, for Heisenberg's uncertainty principle would acquire the status of *certainty;* when $h \to 0$, the fluctuations in the position and momentum will vanish, $\Delta x \to 0$ and $\Delta p \to 0$. Thus, the position and momentum can be measured simultaneously with arbitrary accuracy. This implies that the probabilistic assessments of dynamical quantities by quantum mechanics must give way to exact calculations (these ideas will be discussed further when we study the WKB method in Chapter 9).

So, for those cases where the action variables of a system are too large compared to h (or, equivalently, when the lengths of this system are too large compared to its de Broglie wavelength), quantum mechanics gives the same results as classical mechanics.

In the rest of this text, we will deal with the various applications of the Schrödinger equation. We start, in Chapter 4, with the simple case of one-dimensional systems and later on consider more realistic systems.

3.9 Solved Problems

Problem 3.1
A particle of mass m, which moves freely inside an infinite potential well of length a, has the following initial wave function at $t = 0$:

$$\psi(x,0) = \frac{A}{\sqrt{a}}\sin\left(\frac{\pi x}{a}\right) + \sqrt{\frac{3}{5a}}\sin\left(\frac{3\pi x}{a}\right) + \frac{1}{\sqrt{5a}}\sin\left(\frac{5\pi x}{a}\right),$$

where A is a real constant.

(a) Find A so that $\psi(x,0)$ is normalized.

(b) If measurements of the energy are carried out, what are the values that will be found and what are the corresponding probabilities? Calculate the average energy.

(c) Find the wave function $\psi(x,t)$ at any later time t.

(d) Determine the probability of finding the system at a time t in the state $\varphi(x,t) = \sqrt{2/a}\sin(5\pi x/a)\exp(-iE_5 t/\hbar)$; then determine the probability of finding it in the state $\chi(x,t) = \sqrt{2/a}\sin(2\pi x/a)\exp(-iE_2 t/\hbar)$.

Solution
Since the functions

$$\phi_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi x}{a}\right) \tag{3.149}$$

are orthonormal,

$$\langle\phi_n|\phi_m\rangle = \int_0^a \phi_n^*(x)\phi_m(x)\,dx = \frac{2}{a}\int_0^a \sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi x}{a}\right)dx = \delta_{nm}, \tag{3.150}$$

it is more convenient to write $\psi(x,0)$ in terms of $\phi_n(x)$:

$$\begin{aligned}\psi(x,0) &= \frac{A}{\sqrt{a}}\sin\left(\frac{\pi x}{a}\right) + \sqrt{\frac{3}{5a}}\sin\left(\frac{3\pi x}{a}\right) + \frac{1}{\sqrt{5a}}\sin\left(\frac{5\pi x}{a}\right)\\ &= \frac{A}{\sqrt{2}}\phi_1(x) + \sqrt{\frac{3}{10}}\phi_3(x) + \frac{1}{\sqrt{10}}\phi_5(x).\end{aligned} \tag{3.151}$$

(a) Since $\langle\phi_n|\phi_m\rangle = \delta_{nm}$ the normalization of $\psi(x,0)$ yields

$$1 = \langle\psi|\psi\rangle = \frac{A^2}{2} + \frac{3}{10} + \frac{1}{10}, \tag{3.152}$$

or $A = \sqrt{6/5}$; hence

$$\psi(x,0) = \sqrt{\frac{3}{5}}\phi_1(x) + \sqrt{\frac{3}{10}}\phi_3(x) + \frac{1}{\sqrt{10}}\phi_5(x). \tag{3.153}$$

(b) Since the second derivative of (3.149) is given by $d^2\phi_n(x)/dx^2 = -(n^2\pi^2/a^2)\phi_n(x)$, and since the Hamiltonian of a free particle is $\hat{H} = -(\hbar^2/2m)d^2/dx^2$, the expectation value of $\hat{H}$ with respect to $\phi_n(x)$ is

$$E_n = \langle\phi_n|\hat{H}|\phi_n\rangle = -\frac{\hbar^2}{2m}\int_0^a \phi_n^*(x)\frac{d^2\phi_n(x)}{dx^2}dx = \frac{n^2\pi^2\hbar^2}{2ma^2}. \tag{3.154}$$

If a measurement is carried out on the system, we would obtain $E_n = n^2\pi^2\hbar^2/(2ma^2)$ with a corresponding probability of $P_n(E_n) = |\langle\phi_n|\psi\rangle|^2$. Since the initial wave function (3.153) contains only three eigenstates of $\hat{H}$, $\phi_1(x)$, $\phi_3(x)$, and $\phi_5(x)$, the results of the energy measurements along with the corresponding probabilities are

$$E_1 = \langle\phi_1|\hat{H}|\phi_1\rangle = \frac{\pi^2\hbar^2}{2ma^2}, \qquad P_1(E_1) = |\langle\phi_1|\psi\rangle|^2 = \frac{3}{5}, \tag{3.155}$$

$$E_3 = \langle\phi_3|\hat{H}|\phi_3\rangle = \frac{9\pi^2\hbar^2}{2ma^2}, \qquad P_3(E_3) = |\langle\phi_3|\psi\rangle|^2 = \frac{3}{10}, \tag{3.156}$$

$$E_5 = \langle\phi_5|\hat{H}|\phi_5\rangle = \frac{25\pi^2\hbar^2}{2ma^2}, \qquad P_5(E_5) = |\langle\phi_5|\psi\rangle|^2 = \frac{1}{10}. \tag{3.157}$$

The average energy is

$$E = \sum_n P_n E_n = \frac{3}{5}E_1 + \frac{3}{10}E_3 + \frac{1}{10}E_5 = \frac{29\pi^2\hbar^2}{10ma^2}. \tag{3.158}$$

(c) As the initial state $\psi(x,0)$ is given by (3.153), the wave function $\psi(x,t)$ at any later time t is

$$\psi(x,t) = \sqrt{\frac{3}{5}}\phi_1(x)e^{-iE_1t/\hbar} + \sqrt{\frac{3}{10}}\phi_3(x)e^{-iE_3t/\hbar} + \frac{1}{\sqrt{10}}\phi_5(x)e^{-iE_5t/\hbar}, \tag{3.159}$$

where the expressions of E_n are listed in (3.154) and $\phi_n(x)$ in (3.149).

(d) First, let us express $\varphi(x,t)$ in terms of $\phi_n(x)$:

$$\varphi(x,t) = \sqrt{\frac{2}{a}}\sin\left(\frac{5\pi x}{a}\right)e^{-iE_5t/\hbar} = \phi_5(x)e^{-iE_5t/\hbar}. \tag{3.160}$$

The probability of finding the system at a time t in the state $\varphi(x,t)$ is

$$P = |\langle\varphi|\psi\rangle|^2 = \left|\int_0^a \varphi^*(x,t)\psi(x,t)\,dx\right|^2 = \frac{1}{10}\left|\int_0^a \phi_5^*(x)\phi_5(x)\,dx\right|^2 = \frac{1}{10}, \tag{3.161}$$

since $\langle\varphi|\phi_1\rangle = \langle\varphi|\phi_3\rangle = 0$ and $\langle\varphi|\phi_5\rangle = \exp(iE_5t/\hbar)$.

Similarly, since $\chi(x,t) = \sqrt{2/a}\sin(2\pi x/a)\exp(-iE_2t/\hbar) = \phi_2(x)\exp(-iE_2t/\hbar)$, we can easily show that the probability for finding the system in the state $\chi(x,t)$ is zero:

$$P = |\langle\chi|\psi\rangle|^2 = \left|\int_0^a \chi^*(x,t)\psi(x,t)\,dx\right|^2 = 0, \tag{3.162}$$

since $\langle\chi|\phi_1\rangle = \langle\chi|\phi_3\rangle = \langle\chi|\phi_5\rangle = 0$.

Problem 3.2

A particle of mass m, which moves freely inside an infinite potential well of length a, is initially in the state $\psi(x,0) = \sqrt{3/5a}\sin(3\pi x/a) + (1/\sqrt{5a})\sin(5\pi x/a)$.

(a) Find $\psi(x,t)$ at any later time t.

(b) Calculate the probability density $\rho(x,t)$ and the current density, $\vec{j}(x,t)$.

(c) Verify that the probability is conserved, i.e., $\partial\rho/\partial t + \vec{\nabla}\cdot\vec{j}(x,t) = 0$.

Solution

(a) Since $\psi(x,0)$ can be expressed in terms of $\phi_n(x) = \sqrt{2/a}\sin(n\pi x/a)$ as

$$\psi(x,0) = \sqrt{\frac{3}{5a}}\sin\left(\frac{3\pi x}{a}\right) + \frac{1}{\sqrt{5a}}\sin\left(\frac{5\pi x}{a}\right) = \sqrt{\frac{3}{10}}\phi_3(x) + \frac{1}{\sqrt{10}}\phi_5(x), \tag{3.163}$$

we can write

$$\begin{aligned}\psi(x,t) &= \sqrt{\frac{3}{5a}}\sin\left(\frac{3\pi x}{a}\right)e^{-iE_3t/\hbar} + \frac{1}{\sqrt{5a}}\sin\left(\frac{5\pi x}{a}\right)e^{-iE_5t/\hbar} \\ &= \sqrt{\frac{3}{10}}\phi_3(x)e^{-iE_3t/\hbar} + \frac{1}{\sqrt{10}}\phi_5(x)e^{-iE_5t/\hbar},\end{aligned} \tag{3.164}$$

where the expressions for E_n are listed in (3.154): $E_n = n^2\pi^2\hbar^2/(2ma^2)$.

(b) Since $\rho(x,t) = \psi^*(x,t)\psi(x,t)$, where $\psi(x,t)$ is given by (3.164), we can write

$$\rho(x,t) = \frac{3}{10}\phi_3^2(x) + \frac{\sqrt{3}}{10}\phi_3(x)\phi_5(x)\left[e^{i(E_3-E_5)t/\hbar} + e^{-i(E_3-E_5)t/\hbar}\right] + \frac{1}{10}\phi_5^2(x). \tag{3.165}$$

From (3.154) we have $E_3 - E_5 = 9E_1 - 25E_1 = -16E_1 = -8\pi^2\hbar^2/(ma^2)$. Thus, $\rho(x,t)$ becomes

$$\begin{aligned}\rho(x,t) &= \frac{3}{10}\phi_3^2(x) + \frac{\sqrt{3}}{5}\phi_3(x)\phi_5(x)\cos\left(\frac{16E_1t}{\hbar}\right) + \frac{1}{10}\phi_5^2(x) \\ &= \frac{3}{5a}\sin^2\left(\frac{3\pi x}{a}\right) + \frac{2\sqrt{3}}{5a}\sin\left(\frac{3\pi x}{a}\right)\sin\left(\frac{5\pi x}{a}\right)\cos\left(\frac{16E_1t}{\hbar}\right) \\ &\quad + \frac{1}{5a}\sin^2\left(\frac{5\pi x}{a}\right).\end{aligned} \tag{3.166}$$

Since the system is one-dimensional, the action of the gradient operator on $\psi(x,t)$ and $\psi^*(x,t)$ is given by $\vec{\nabla}\psi(x,t) = (d\psi(x,t)/dx)\vec{i}$ and $\vec{\nabla}\psi^*(x,t) = (d\psi^*(x,t)/dx)\vec{i}$. We can thus write the current density $\vec{J}(x,t) = (i\hbar/2m)\left(\psi(x,t)\vec{\nabla}\psi^*(x,t) - \psi^*(x,t)\vec{\nabla}\psi(x,t)\right)$ as

$$\vec{J}(x,t) = \frac{i\hbar}{2m}\left(\psi(x,t)\frac{d\psi^*(x,t)}{dx} - \psi^*(x,t)\frac{d\psi(x,t)}{dx}\right)\vec{i}. \tag{3.167}$$

Using (3.164) we have

$$\frac{d\psi(x,t)}{dx} = \frac{3\pi}{a}\sqrt{\frac{3}{5a}}\cos\left(\frac{3\pi x}{a}\right)e^{-iE_3t/\hbar} + \frac{5\pi}{a}\frac{1}{\sqrt{5a}}\cos\left(\frac{5\pi x}{a}\right)e^{-iE_5t/\hbar}, \tag{3.168}$$

$$\frac{d\psi^*(x,t)}{dx} = \frac{3\pi}{a}\sqrt{\frac{3}{5a}}\cos\left(\frac{3\pi x}{a}\right)e^{iE_3t/\hbar} + \frac{5\pi}{a}\frac{1}{\sqrt{5a}}\cos\left(\frac{5\pi x}{a}\right)e^{iE_5t/\hbar}. \tag{3.169}$$

A straightforward calculation yields

$$\begin{aligned}\psi\frac{d\psi^*}{dx} - \psi^*\frac{d\psi}{dx} = -2i\pi\frac{\sqrt{3}}{5a^2}&\left[5\sin\left(\frac{3\pi x}{a}\right)\cos\left(\frac{5\pi x}{a}\right) - 3\sin\left(\frac{5\pi x}{a}\right)\cos\left(\frac{3\pi x}{a}\right)\right] \\ &\times \sin\left(\frac{E_3 - E_5}{\hbar}t\right).\end{aligned} \tag{3.170}$$

Inserting this into (3.167) and using $E_3 - E_5 = -16E_1$, we have

$$\vec{j}(x,t) = -\frac{\pi\hbar}{m}\frac{\sqrt{3}}{5a^2}\left[5\sin\left(\frac{3\pi x}{a}\right)\cos\left(\frac{5\pi x}{a}\right) - 3\sin\left(\frac{5\pi x}{a}\right)\cos\left(\frac{3\pi x}{a}\right)\right]\sin\left(\frac{16E_1 t}{\hbar}\right)\vec{i}. \tag{3.171}$$

(c) Performing the time derivative of (3.166) and using the expression $32\sqrt{3}E_1/(5a\hbar) = 16\pi^2\hbar\sqrt{3}/(5ma^3)$, since $E_1 = \pi^2\hbar^2/(2ma^2)$, we obtain

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= -\frac{32\sqrt{3}E_1}{5a\hbar}\sin\left(\frac{3\pi x}{a}\right)\sin\left(\frac{5\pi x}{a}\right)\sin\left(\frac{16E_1 t}{\hbar}\right)\\ &= -\frac{16\pi^2\hbar\sqrt{3}}{5ma^3}\sin\left(\frac{3\pi x}{a}\right)\sin\left(\frac{5\pi x}{a}\right)\sin\left(\frac{16E_1 t}{\hbar}\right),\end{aligned} \tag{3.172}$$

Now, taking the divergence of (3.171), we end up with

$$\vec{\nabla}\cdot\vec{j}(x,t) = \frac{dJ(x,t)}{dx} = \frac{16\pi^2\hbar\sqrt{3}}{5ma^3}\sin\left(\frac{3\pi x}{a}\right)\sin\left(\frac{5\pi x}{a}\right)\sin\left(\frac{16E_1 t}{\hbar}\right). \tag{3.173}$$

The addition of (3.172) and (3.173) confirms the conservation of probability:

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{j}(x,t) = 0. \tag{3.174}$$

Problem 3.3

Consider a one-dimensional particle which is confined within the region $0 \le x \le a$ and whose wave function is $\Psi(x,t) = \sin(\pi x/a)\exp(-i\omega t)$.

(a) Find the potential $V(x)$.

(b) Calculate the probability of finding the particle in the interval $a/4 \le x \le 3a/4$.

Solution

(a) Since the first time derivative and the second x derivative of $\Psi(x,t)$ are given by $\partial\Psi(x,t)/\partial t = -i\omega\Psi(x,t)$ and $\partial^2\Psi(x,t)/\partial x^2 = -(\pi^2/a^2)\Psi(x,t)$, the Schrödinger equation (3.73) yields

$$i\hbar(-i\omega)\Psi(x,t) = \frac{\hbar^2}{2m}\frac{\pi^2}{a^2}\Psi(x,t) + \hat{V}(x,t)\Psi(x,t). \tag{3.175}$$

Hence $V(x,t)$ is time independent and given by $V(x) = \hbar\omega - \hbar^2\pi^2/(2ma^2)$.

(b) The probability of finding the particle in the interval $a/4 \le x \le 3a/4$ can be obtained from (3.8):

$$P = \frac{\int_{a/4}^{3a/4}|\psi(x)|^2dx}{\int_0^a|\psi(x)|^2dx} = \frac{\int_{a/4}^{3a/4}\sin^2(\pi x/a)\,dx}{\int_0^a\sin^2(\pi x/a)\,dx} = \frac{2+\pi}{2\pi} = 0.82 \tag{3.176}$$

Problem 3.4

A system is initially in the state $|\psi_0\rangle = [\sqrt{2}|\phi_1\rangle + \sqrt{3}|\phi_2\rangle + |\phi_3\rangle + |\phi_4\rangle]/\sqrt{7}$, where $|\phi_n\rangle$ are eigenstates of the system's Hamiltonian such that $\hat{H}|\phi_n\rangle = n^2\mathcal{E}_0|\phi_n\rangle$.

(a) If energy is measured, what values will be obtained and with what probabilities?

(b) Consider an operator $\hat{A}$ whose action on $|\phi_n\rangle$ is defined by $\hat{A}|\phi_n\rangle = (n+1)a_0|\phi_n\rangle$. If A is measured, what values will be obtained and with what probabilities?

(c) Suppose that a measurement of the energy yields $4\mathcal{E}_0$. If we measure A immediately afterwards, what value will be obtained?

Solution

(a) A measurement of the energy yields $E_n = \langle\phi_n|\hat{H}|\phi_n\rangle = n^2\mathcal{E}_0$, that is

$$E_1 = \mathcal{E}_0, \quad E_2 = 4\mathcal{E}_0, \quad E_3 = 9\mathcal{E}_0, \quad E_4 = 16\mathcal{E}_0. \tag{3.177}$$

Since $|\psi_0\rangle$ is normalized, $\langle\psi_0 \mid \psi_0\rangle = (2 + 3 + 1 + 1)/7 = 1$, and using (3.3), we can write the probabilities corresponding to (3.177) as $P(E_n) = \left|\langle\phi_n|\psi_0\rangle\right|^2 / \langle\psi_0 \mid \psi_0\rangle = \left|\langle\phi_n|\psi_0\rangle\right|^2$; hence, using the fact that $\langle\phi_n|\phi_m\rangle = \delta_{nm}$, we have

$$P(E_1) = \left|\sqrt{\frac{2}{7}}\langle\phi_1|\phi_1\rangle\right|^2 = \frac{2}{7}, \qquad P(E_2) = \left|\sqrt{\frac{3}{7}}\langle\phi_2|\phi_2\rangle\right|^2 = \frac{3}{7}, \tag{3.178}$$

$$P(E_3) = \left|\frac{1}{\sqrt{7}}\langle\phi_3|\phi_3\rangle\right|^2 = \frac{1}{7}, \qquad P(E_4) = \left|\frac{1}{\sqrt{7}}\langle\phi_4|\phi_4\rangle\right|^2 = \frac{1}{7}. \tag{3.179}$$

(b) Similarly, a measurement of the observable $\hat{A}$ yields $a_n = \langle\phi_n|\hat{A}|\phi_n\rangle = (n + 1)a_0$; that is,

$$a_1 = 2a_0, \quad a_2 = 3a_0, \quad a_3 = 4a_0, \quad a_4 = 5a_0. \tag{3.180}$$

Again, using (3.3) and since $|\psi_0\rangle$ is normalized, we can ascertain that the probabilities corresponding to the values (3.180) are given by $P(a_n) = \left|\langle\phi_n|\psi_0\rangle\right|^2 / \langle\psi_0 \mid \psi_0\rangle = \left|\langle\phi_n|\psi_0\rangle\right|^2$, or

$$P(a_1) = \left|\sqrt{\frac{2}{7}}\langle\phi_1|\phi_1\rangle\right|^2 = \frac{2}{7}, \qquad P(a_2) = \left|\sqrt{\frac{3}{7}}\langle\phi_2|\phi_2\rangle\right|^2 = \frac{3}{7}, \tag{3.181}$$

$$P(a_3) = \left|\frac{1}{\sqrt{7}}\langle\phi_3|\phi_3\rangle\right|^2 = \frac{1}{7}, \qquad P(a_4) = \left|\frac{1}{\sqrt{7}}\langle\phi_4|\phi_4\rangle\right|^2 = \frac{1}{7}. \tag{3.182}$$

(c) An energy measurement that yields $4\mathcal{E}_0$ implies that the system is left in the state $|\phi_2\rangle$. A measurement of the observable A immediately afterwards leads to

$$\langle\phi_2|\hat{A}|\phi_2\rangle = 3a_0\langle\phi_2|\phi_2\rangle = 3a_0. \tag{3.183}$$

Problem 3.5

(a) Assuming that the system of Problem 3.4 is initially in the state $|\phi_3\rangle$, what values for the energy and the observable A will be obtained if we measure: (i)H first then A, (ii) A first then H?

(b) Compare the results obtained in (i) and (ii) and infer whether $\hat{H}$ and $\hat{A}$ are compatible. Calculate $[\hat{H}, \hat{A}]|\phi_3\rangle$.

(c) Consider now an other operator $\hat{B}$ whose action on $|\phi_n\rangle$ is defined by $\hat{B}|\phi_n\rangle = nb_0|\phi_{n+1}\rangle$. Repeat Questions (a) and (b) for the operator $\hat{B}$.

Solution

(a) (i) The measurement of H first then A is represented by $\hat{A}\hat{H}|\phi_3\rangle$. Using the relations $\hat{H}|\phi_n\rangle = n^2\mathcal{E}_0|\phi_n\rangle$ and $\hat{A}|\phi_n\rangle = (n + 1)a_0|\phi_n\rangle$, we have

$$\hat{A}\hat{H}|\phi_3\rangle = 9\mathcal{E}_0\hat{A}|\phi_3\rangle = 36\,\mathcal{E}_0a_0|\phi_3\rangle. \tag{3.184}$$

(ii) Measuring A first and then H, we will obtain

$$\hat{H}\hat{A}|\phi_3\rangle = 4a_0\hat{H}|\phi_3\rangle = 36\,\mathcal{E}_0 a_0|\phi_3\rangle. \tag{3.185}$$

(b) Equations (3.184) and (3.185) show that the actions of $\hat{A}\hat{H}$ and $\hat{H}\hat{A}$ yield the same result. This means that $\hat{H}$ and $\hat{A}$ commute; hence they are compatible. We can thus write

$$[\hat{H},\ \hat{A}]|\phi_3\rangle = (36 - 36)\mathcal{E}_0 a_0|\phi_3\rangle = 0. \tag{3.186}$$

(c) (i) The measurement of H first then B is represented by $\hat{B}\hat{H}|\phi_3\rangle$. Using the relations $\hat{H}|\phi_n\rangle = n^2\mathcal{E}_0|\phi_n\rangle$ and $\hat{B}|\phi_n\rangle = nb_0|\phi_{n+1}\rangle$, we have

$$\hat{B}\hat{H}|\phi_3\rangle = 9\mathcal{E}_0\hat{B}|\phi_3\rangle = 27\,\mathcal{E}_0 b_0|\phi_4\rangle. \tag{3.187}$$

(ii) Measuring B first and then H, we will obtain

$$\hat{H}\hat{B}|\phi_3\rangle = 3b_0\hat{H}|\phi_4\rangle = 48\,\mathcal{E}_0 b_0|\phi_4\rangle. \tag{3.188}$$

Equations (3.187) and (3.188) show that the actions of $\hat{B}\hat{H}$ and $\hat{H}\hat{B}$ yield different results. This means that $\hat{H}$ and $\hat{B}$ do not commute; hence they are not compatible. We can thus write

$$[\hat{H},\ \hat{B}]|\phi_3\rangle = (48 - 27)\,\mathcal{E}_0 b_0|\phi_4\rangle = 21\,\mathcal{E}_0 b_0|\phi_4\rangle. \tag{3.189}$$

Problem 3.6

Consider a physical system whose Hamiltonian H and initial state $|\psi_0\rangle$ are given by

$$H = \mathcal{E}\begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad |\psi_0\rangle = \frac{1}{\sqrt{5}}\begin{pmatrix} 1-i \\ 1-i \\ 1 \end{pmatrix},$$

where $\mathcal{E}$ has the dimensions of energy.

(a) What values will we obtain when measuring the energy and with what probabilities?

(b) Calculate $\langle\hat{H}\rangle$, the expectation value of the Hamiltonian.

Solution

(a) The results of the energy measurement are given by the eigenvalues of H. A diagonalization of H yields a nondegenerate eigenenergy $E_1 = \mathcal{E}$ and a doubly degenerate value $E_2 = E_3 = -\mathcal{E}$ whose respective eigenvectors are given by

$$|\phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \qquad |\phi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix}, \qquad |\phi_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; \tag{3.190}$$

these eigenvectors are orthogonal since H is Hermitian. Note that the initial state $|\psi_0\rangle$ can be written in terms of $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ as follows:

$$|\psi_0\rangle = \frac{1}{\sqrt{5}}\begin{pmatrix} 1-i \\ 1-i \\ 1 \end{pmatrix} = \sqrt{\frac{2}{5}}|\phi_1\rangle + \sqrt{\frac{2}{5}}|\phi_2\rangle + \frac{1}{\sqrt{5}}|\phi_3\rangle. \tag{3.191}$$

Since $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ are orthonormal, the probability of measuring $E_1 = \mathcal{E}$ is given by

$$P_1(E_1) = \left|\langle\phi_1|\psi_0\rangle\right|^2 = \left|\sqrt{\frac{2}{5}}\langle\phi_1|\phi_1\rangle\right|^2 = \frac{2}{5}. \tag{3.192}$$

Now, since the other eigenvalue is doubly degenerate, $E_2 = E_3 = -\mathcal{E}$, the probability of measuring $-\mathcal{E}$ can be obtained from (3.5):

$$P_2(E_2) = \left|\langle\phi_2|\psi_0\rangle\right|^2 + \left|\langle\phi_3|\psi_0\rangle\right|^2 = \frac{2}{5} + \frac{1}{5} = \frac{3}{5}. \tag{3.193}$$

(b) From (3.192) and (3.193), we have

$$\langle\hat{H}\rangle = P_1E_1 + P_2E_2 = \frac{2}{5}\mathcal{E} - \frac{3}{5}\mathcal{E} = -\frac{1}{5}\mathcal{E}. \tag{3.194}$$

We can obtain the same result by calculating the expectation value of $\hat{H}$ with respect to $|\psi_0\rangle$. Since $\langle\psi_0|\psi_0\rangle = 1$, we have $\langle\hat{H}\rangle = \langle\psi_0|\hat{H}|\psi_0\rangle/\langle\psi_0|\psi_0\rangle = \langle\psi_0|\hat{H}|\psi_0\rangle$:

$$\langle\hat{H}\rangle = \langle\psi_0|\hat{H}|\psi_0\rangle = \frac{\mathcal{E}}{5}\begin{pmatrix} 1+i & 1+i & 1 \end{pmatrix}\begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1-i \\ 1-i \\ 1 \end{pmatrix} = -\frac{1}{5}\mathcal{E}. \tag{3.195}$$

Problem 3.7
Consider a system whose Hamiltonian H and an operator A are given by the matrices

$$H = \mathcal{E}_0\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad A = a\begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

where $\mathcal{E}_0$ has the dimensions of energy.

(a) If we measure the energy, what values will we obtain?

(b) Suppose that when we measure the energy, we obtain a value of $-\mathcal{E}_0$. Immediately afterwards, we measure A. What values will we obtain for A and what are the probabilities corresponding to each value?

(c) Calculate the uncertainty ΔA.

Solution

(a) The possible energies are given by the eigenvalues of H. A diagonalization of H yields three nondegenerate eigenenergies $E_1 = 0$, $E_2 = -\mathcal{E}_0$, and $E_3 = 2\mathcal{E}_0$. The respective eigenvectors are

$$|\phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad |\phi_2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad |\phi_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}; \tag{3.196}$$

these eigenvectors are orthonormal.

(b) If a measurement of the energy yields $-\mathcal{E}_0$, this means that the system is left in the state $|\phi_2\rangle$. When we measure the next observable, A, the system is in the state $|\phi_2\rangle$. The result we obtain for A is given by any of the eigenvalues of A. A diagonalization of A yields three nondegenerate values: $a_1 = -\sqrt{17}a$, $a_2 = 0$, and $a_3 = \sqrt{17}a$; their respective eigenvectors are given by

$$|a_1\rangle = \frac{1}{\sqrt{34}}\begin{pmatrix} 4 \\ -\sqrt{17} \\ 1 \end{pmatrix}, \qquad |a_2\rangle = \frac{1}{\sqrt{17}}\begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix}, \qquad |a_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 4 \\ \sqrt{17} \\ 1 \end{pmatrix}. \tag{3.197}$$

Thus, when measuring A on a system which is in the state $|\phi_2\rangle$, the probability of finding $-\sqrt{17}a$ is given by

$$P_1(a_1) = |\langle a_1|\phi_2\rangle|^2 = \left|\frac{1}{\sqrt{34}}\begin{pmatrix} 4 & -\sqrt{17} & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right|^2 = \frac{1}{34}. \tag{3.198}$$

Similarly, the probabilities of measuring 0 and $\sqrt{17}a$ are

$$P_2(a_2) = |\langle a_2|\phi_2\rangle|^2 = \left|\frac{1}{\sqrt{17}}\begin{pmatrix} 1 & 0 & -4 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right|^2 = \frac{16}{17}, \tag{3.199}$$

$$P_3(a_3) = |\langle a_3|\phi_2\rangle|^2 = \left|\frac{1}{\sqrt{34}}\begin{pmatrix} 4 & \sqrt{17} & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right|^2 = \frac{1}{34}. \tag{3.200}$$

(c) Since the system, when measuring A is in the state $|\phi_2\rangle$, the uncertainty ΔA is given by $\Delta A = \sqrt{\langle\phi_2|A^2|\phi_2\rangle - \langle\phi_2|A|\phi_2\rangle^2}$, where

$$\langle\phi_2|A|\phi_2\rangle = a\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0, \tag{3.201}$$

$$\langle\phi_2|A^2|\phi_2\rangle = a^2\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = a^2. \tag{3.202}$$

Thus we have $\Delta A = a$.

Problem 3.8
Consider a system whose state and two observables are given by

$$|\psi(t)\rangle = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \qquad A = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(a) What is the probability that a measurement of A at time t yields -1?

(b) Let us carry out a set of two measurements where B is measured first and then, immediately afterwards, A is measured. Find the probability of obtaining a value of 0 for B and a value of 1 for A.

(c) Now we measure A first then, immediately afterwards, B. Find the probability of obtaining a value of 1 for A and a value of 0 for B.

(d) Compare the results of (b) and (c). Explain.

(e) Which among the sets of operators $\{\hat{A}\}$, $\{\hat{B}\}$, and $\{\hat{A}, \hat{B}\}$ form a complete set of commuting operators (CSCO)?

Solution

(a) A measurement of A yields any of the eigenvalues of A which are given by $a_1 = -1$, $a_2 = 0$, $a_3 = 1$; the respective (normalized) eigenstates are

$$|a_1\rangle = \frac{1}{2}\begin{pmatrix} -1 \\ \sqrt{2} \\ -1 \end{pmatrix}, \qquad |a_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \qquad |a_3\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}. \tag{3.203}$$

The probability of obtaining $a_1 = -1$ is

$$P(-1) = \frac{|\langle a_1|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} = \frac{1}{6}\left|\frac{1}{2}\begin{pmatrix} -1 & \sqrt{2} & -1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}\right|^2 = \frac{1}{3}, \tag{3.204}$$

where we have used the fact that $\langle\psi(t)|\psi(t)\rangle = \begin{pmatrix} -1 & 2 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = 6$.

(b) A measurement of B yields a value which is equal to any of the eigenvalues of B: $b_1 = -1$, $b_2 = 0$, and $b_3 = 1$; their corresponding eigenvectors are

$$|b_1\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad |b_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |b_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{3.205}$$

Since the system was in the state $|\psi(t)\rangle$, the probability of obtaining the value $b_2 = 0$ for B is

$$P(b_2) = \frac{|\langle b_2|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} = \frac{1}{6}\left|\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}\right|^2 = \frac{2}{3}. \tag{3.206}$$

We deal now with the measurement of the other observable, A. The observables A and B do not have common eigenstates, since they do not commute. After measuring B (the result is $b_2 = 0$), the system is left, according to Postulate 3, in a state $|\phi\rangle$ which can be found by projecting $|\psi(t)\rangle$ onto $|b_2\rangle$:

$$|\phi\rangle = |b_2\rangle\langle b_2|\psi(t)\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}. \tag{3.207}$$

The probability of finding 1 when we measure A is given by

$$P(a_3) = \frac{|\langle a_3|\phi\rangle|^2}{\langle\phi|\phi\rangle} = \frac{1}{4}\left|\frac{1}{2}\begin{pmatrix} 1 & \sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}\right|^2 = \frac{1}{2}, \tag{3.208}$$

since $\langle\phi|\phi\rangle = 4$. In summary, when measuring B then A, the probability of finding a value of 0 for B and 1 for A is given by the product of the probabilities (3.206) and (3.208):

$$P(b_2, a_3) = P(b_2)P(a_3) = \frac{2}{3}\frac{1}{2} = \frac{1}{3}. \tag{3.209}$$

(c) Next we measure A first then B. Since the system is in the state $|\psi(t)\rangle$, the probability of measuring $a_3 = 1$ for A is given by

$$P'(a_3) = \frac{|\langle a_3|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} = \frac{1}{6}\left|\frac{1}{2}\begin{pmatrix} 1 & \sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}\right|^2 = \frac{1}{3}, \tag{3.210}$$

where we have used the expression (3.203) for $|a_3\rangle$.

We then proceed to the measurement of B. The state of the system just after measuring A (with a value $a_3 = 1$) is given by a projection of $|\psi(t)\rangle$ onto $|a_3\rangle$:

$$|\chi\rangle = |a_3\rangle\langle a_3|\psi(t)\rangle = \frac{1}{4}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}\begin{pmatrix} 1 & \sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}. \tag{3.211}$$

So the probability of finding a value of $b_2 = 0$ when measuring B is given by

$$P'(b_2) = \frac{|\langle b_2|\chi\rangle|^2}{\langle\chi|\chi\rangle} = \frac{1}{2}\left|\frac{\sqrt{2}}{2}\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}\right|^2 = \frac{1}{2}, \tag{3.212}$$

since $\langle\chi|\chi\rangle = 2$.

So when measuring A then B, the probability of finding a value of 1 for A and 0 for B is given by the product of the probabilities (3.212) and (3.210):

$$P(a_3, b_2) = P'(a_3)P'(b_2) = \frac{1}{3}\frac{1}{2} = \frac{1}{6}. \tag{3.213}$$

(d) The probabilities $P(b_2, a_3)$ and $P(a_3, b_2)$, as shown in (3.209) and (3.213), are different. This is expected, since A and B do not commute. The result of the successive measurements of A and B therefore depends on the order in which they are carried out. The probability of obtaining 0 for B then 1 for A is equal to $\frac{1}{3}$. On the other hand, the probability of obtaining 1 for A then 0 for B is equal to $\frac{1}{6}$. However, if the observables A and B commute, the result of the measurements will not depend on the order in which they are carried out (this idea is illustrated in the following solved problem).

(e) As stated in the text, any operator with non-degenerate eigenvalues constitutes, all by itself, a CSCO. Hence each of $\{\hat{A}\}$ and $\{\hat{B}\}$ forms a CSCO, since their eigenvalues are not degenerate. However, the set $\{\hat{A}, \hat{B}\}$ does not form a CSCO since the opertators $\{\hat{A}\}$ and $\{\hat{B}\}$ do not commute.

Problem 3.9
Consider a system whose state and two observables A and B are given by

$$|\psi(t)\rangle = \frac{1}{6}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}, \qquad A = \frac{1}{\sqrt{2}}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & i \\ 0 & -i & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}.$$

(a) We perform a measurement where A is measured first and then, immediately afterwards, B is measured. Find the probability of obtaining a value of 0 for A and a value of 1 for B.

(b) Now we measure B first then, immediately afterwards, A. Find the probability of obtaining a value of 1 for B and a value of 0 for A.

(c) Compare the results of (b) and (c). Explain.

(d) Which among the sets of operators $\{\hat{A}\}$, $\{\hat{B}\}$, and $\{\hat{A}, \hat{B}\}$ form a complete set of commuting operators (CSCO)?

Solution

(a) A measurement of A yields any of the eigenvalues of A which are given by $a_1 = 0$ (not degenerate) and $a_2 = a_3 = 2$ (doubly degenerate); the respective (normalized) eigenstates are

$$|a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}, \qquad |a_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}, \qquad |a_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{3.214}$$

The probability that a measurement of A yields $a_1 = 0$ is given by

$$P(a_1) = \frac{|\langle a_1|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} = \frac{36}{17}\left|\frac{1}{\sqrt{2}}\frac{1}{6}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}\right|^2 = \frac{8}{17}, \tag{3.215}$$

where we have used the fact that $\langle\psi(t)|\psi(t)\rangle = \frac{1}{36}\begin{pmatrix} 1 & 0 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} = \frac{17}{36}$.

Since the system was initially in the state $|\psi(t)\rangle$, after a measurement of A yields $a_1 = 0$, the system is left, as mentioned in Postulate 3, in the following state:

$$|\phi\rangle = |a_1\rangle\langle a_1|\psi(t)\rangle = \frac{1}{2}\frac{1}{6}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}. \tag{3.216}$$

As for the measurement of B, we obtain any of the eigenvalues $b_1 = -1$, $b_2 = b_3 = 1$; their corresponding eigenvectors are

$$|b_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}, \qquad |b_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}, \qquad |b_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{3.217}$$

Since the system is now in the state $|\phi\rangle$, the probability of obtaining the (doubly degenerate) value $b_2 = b_3 = 1$ for B is

$$\begin{aligned} P(b_2) &= \frac{|\langle b_2|\phi\rangle|^2}{\langle\phi|\phi\rangle} + \frac{|\langle b_3|\phi\rangle|^2}{\langle\phi|\phi\rangle} \\ &= \frac{1}{2}\left|\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}\right|^2 + \frac{1}{2}\left|\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}\right|^2 \\ &= 1. \end{aligned} \tag{3.218}$$

The reason $P(b_2) = 1$ is because the new state $|\phi\rangle$ is an eigenstate of B; in fact $|\phi\rangle = \sqrt{2}/3|b_2\rangle$.

In sum, when measuring A then B, the probability of finding a value of 0 for A and 1 for B is given by the product of the probabilities (3.215) and (3.218):

$$P(a_1, b_2) = P(a_1)P(b_2) = \frac{8}{17}. \tag{3.219}$$

(b) Next we measure B first then A. Since the system is in the state $|\psi(t)\rangle$ and since the value $b_2 = b_3 = 1$ is doubly degenerate, the probability of measuring 1 for B is given by

$$\begin{aligned} P'(b_2) &= \frac{|\langle b_2|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} + \frac{|\langle b_3|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} \\ &= \frac{36}{17}\frac{1}{36}\left[\left|\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}\right|^2 + \left|\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}\right|^2\right] \\ &= \frac{9}{17}. \end{aligned} \tag{3.220}$$

We now proceed to the measurement of A. The state of the system immediately after measuring B (with a value $b_2 = b_3 = 1$) is given by a projection of $|\psi(t)\rangle$ onto $|b_2\rangle$, and $|b_3\rangle$

$$\begin{aligned} |\chi\rangle &= |b_2\rangle\langle b_2|\psi(t)\rangle + |b_3\rangle\langle b_3|\psi(t)\rangle \\ &= \frac{1}{12}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} \\ &= \frac{1}{6}\begin{pmatrix} 1 \\ -2i \\ 2i \end{pmatrix}. \end{aligned} \tag{3.221}$$

So the probability of finding a value of $a_1 = 0$ when measuring A is given by

$$P'(a_1) = \frac{|\langle a_1|\chi\rangle|^2}{\langle\chi|\chi\rangle} = \frac{36}{9}\left|\frac{1}{6\sqrt{2}}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -2i \\ 2i \end{pmatrix}\right|^2 = \frac{8}{9}, \tag{3.222}$$

since $\langle\chi|\chi\rangle = \frac{9}{36}$.

Therefore, when measuring B then A, the probability of finding a value of 1 for B and 0 for A is given by the product of the probabilities (3.220) and (3.222):

$$P(b_2, a_3) = P'(b_2)P'(a_1) = \frac{9}{17}\frac{8}{9} = \frac{8}{17}. \tag{3.223}$$

(c) The probabilities $P(a_1, b_2)$ and $P(b_2, a_1)$, as shown in (3.219) and (3.223), are equal. This is expected since A and B do commute. The result of the successive measurements of A and B does not depend on the order in which they are carried out.

(d) Neither $\{\hat{A}\}$ nor $\{\hat{B}\}$ forms a CSCO since their eigenvalues are degenerate. The set $\{\hat{A}, \hat{B}\}$, however, does form a CSCO since the opertators $\{\hat{A}\}$ and $\{\hat{B}\}$ commute. The set of eigenstates that are common to $\{\hat{A}, \hat{B}\}$ are given by

$$|a_2, b_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}, \qquad |a_1, b_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}, \qquad |a_3, b_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{3.224}$$

Problem 3.10

Consider a physical system which has a number of observables that are represented by the following matrices:

$$A = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 3 \\ 0 & 3 & 0 \end{pmatrix}, \; C = \begin{pmatrix} 0 & 3 & 0 \\ 3 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix}, \; D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}.$$

(a) Find the results of the measurements of these observables.

(b) Which among these observables are compatible? Give a basis of eigenvectors common to these observables.

(c) Which among the sets of operators $\{\hat{A}\}$, $\{\hat{B}\}$, $\{\hat{C}\}$, $\{\hat{D}\}$ and their various combinations, such as $\{\hat{A}, \hat{B}\}$, $\{\hat{A}, \hat{C}\}$, $\{\hat{B}, \hat{C}\}$, $\{\hat{A}, \hat{D}\}$, $\{\hat{A}, \hat{B}, \hat{C}\}$, form a complete set of commuting operators (CSCO)?

Solution

(a) The measurements of A, B, C and D yield $a_1 = -1, a_2 = 3, a_3 = 5, b_1 = -3, b_2 = 1, b_3 = 3$, $c_1 = -\sqrt{13}, c_2 = 0, c_3 = \sqrt{13}, d_1 = -1, d_2 = d_3 = 1$; the respective eigenvectors of A, B, C and D are

$$|a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \qquad |a_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad |a_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{3.225}$$

$$|b_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \qquad |b_2\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |b_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \tag{3.226}$$

$$|c_1\rangle = \frac{1}{\sqrt{26}}\begin{pmatrix} 3 \\ -\sqrt{13} \\ 2 \end{pmatrix}, \; |c_2\rangle = \frac{1}{\sqrt{13}}\begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}, \; |c_3\rangle = \frac{1}{\sqrt{26}}\begin{pmatrix} 3 \\ \sqrt{13} \\ 2 \end{pmatrix}, \tag{3.227}$$

$$|d_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}, \qquad |d_2\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |d_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix}. \tag{3.228}$$

(b) We can verify that, among the observables A, B, C, and D, only A and B are compatible, since the matrices A and B commute; the rest do not commute with one another (neither A nor B commutes with C or D; C and D do not commute).

From (3.225) and (3.226) we see that the three states $|a_1, b_1\rangle$, $|a_2, b_3\rangle$, $|a_3, b_2\rangle$,

$$|a_1, b_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \qquad |a_2, b_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad |a_3, b_2\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{3.229}$$

form a common, complete basis for A and B, since $\hat{A}|a_n, b_m\rangle = a_n|a_n, b_m\rangle$ and $\hat{B}|a_n, b_m\rangle = b_m|a_n, b_m\rangle$.

(c) First, since the eigenvalues of the operators $\{\hat{A}\}$, $\{\hat{B}\}$, and $\{\hat{C}\}$ are all nondegenerate, each one of $\{\hat{A}\}$, $\{\hat{B}\}$, and $\{\hat{C}\}$ forms separately a CSCO. Additionally, since two eigenvalues of $\{\hat{D}\}$ are degenerate ($d_2 = d_3 = 1$), the operator $\{\hat{D}\}$ does not form a CSCO.

Now, among the various combinations $\{\hat{A}, \hat{B}\}$, $\{\hat{A}, \hat{C}\}$, $\{\hat{B}, \hat{C}\}$, $\{\hat{A}, \hat{D}\}$, and $\{\hat{A}, \hat{B}, \hat{C}\}$, only $\{\hat{A}, \hat{B}\}$ forms a CSCO, because $\{\hat{A}\}$ and $\{\hat{B}\}$ are the only operators that commute; the set of their joint eigenvectors are given by $|a_1, b_1\rangle$, $|a_2, b_3\rangle$, $|a_3, b_2\rangle$.

Problem 3.11

Consider a system whose initial state $|\psi(0)\rangle$ and Hamiltonian are given by

$$|\psi(0)\rangle = \frac{1}{5}\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}, \qquad H = \epsilon\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 5 & 0 \end{pmatrix},$$

where ϵ has the dimensions of an energy.

(a) If a measurement of the energy is carried out, what values would we obtain and with what probabilities?

(b) Find the state of the system at a later time t; you may need to expand $|\psi(0)\rangle$ in terms of the eigenvectors of H.

(c) Find the total energy of the system at time $t = 0$ and any later time t; are these values different?

(d) Does $\{\hat{H}\}$ form a complete set of commuting operators?

Solution

(a) A measurement of the energy yields the values $E_1 = -5\epsilon$, $E_2 = 3\epsilon$, $E_3 = 5\epsilon$; the respective (orthonormal) eigenvectors of these values are

$$|\phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \qquad |\phi_2\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |\phi_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}. \tag{3.230}$$

The probabilities of finding the values $E_1 = -5$, $E_2 = 3$, $E_3 = 5$ are given by

$$P(E_1) = \left|\langle\phi_1|\psi(0)\rangle\right|^2 = \left|\frac{1}{5\sqrt{2}}\begin{pmatrix} 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}\right|^2 = \frac{8}{25}, \tag{3.231}$$

$$P(E_2) = \left|\langle\phi_2|\psi(0)\rangle\right|^2 = \left|\frac{1}{5}\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}\right|^2 = \frac{9}{25}, \tag{3.232}$$

$$P(E_3) = \left|\langle\phi_3|\psi(0)\rangle\right|^2 = \left|\frac{1}{5\sqrt{2}}\begin{pmatrix} 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}\right|^2 = \frac{8}{25}. \tag{3.233}$$

(b) To find $|\psi(t)\rangle$ we need to expand $|\psi(0)\rangle$ in terms of the eigenvectors (3.230):

$$|\psi(0)\rangle = \frac{1}{5}\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} = \frac{2\sqrt{2}}{5}|\phi_1\rangle + \frac{3}{5}|\phi_2\rangle + \frac{2\sqrt{2}}{5}|\phi_3\rangle; \tag{3.234}$$

hence

$$|\psi(t)\rangle = \frac{2\sqrt{2}}{5}e^{-iE_1t/\hbar}|\phi_1\rangle + \frac{3}{5}e^{-iE_2t/\hbar}|\phi_2\rangle + \frac{2\sqrt{2}}{5}e^{-iE_3t/\hbar}|\phi_3\rangle = \frac{1}{5}\begin{pmatrix} 3e^{-3i\epsilon t/\hbar} \\ -4i\sin(5\epsilon t/\hbar) \\ 4\cos(5\epsilon t/\hbar) \end{pmatrix}. \tag{3.235}$$

(c) We can calculate the energy at time $t = 0$ in three quite different ways. The first method uses the bra-ket notation. Since $\langle\psi(0)|\psi(0)\rangle = 1$, $\langle\phi_n|\phi_m\rangle = \delta_{nm}$ and since $\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$, we have

$$\begin{aligned} E(0) = \langle\psi(0)|\hat{H}|\psi(0)\rangle &= \frac{8}{25}\langle\phi_1|\hat{H}|\phi_1\rangle + \frac{9}{25}\langle\phi_2|\hat{H}|\phi_2\rangle + \frac{8}{25}\langle\phi_3|\hat{H}|\phi_3\rangle \\ &= \frac{8}{25}(-5)\epsilon + \frac{9}{25}(3)\epsilon + \frac{8}{25}(5)\epsilon = \frac{27}{25}\epsilon. \end{aligned} \tag{3.236}$$

The second method uses matrix algebra:

$$E(0) = \langle\psi(0)|\hat{H}|\psi(0)\rangle = \frac{\epsilon}{25}\begin{pmatrix} 3 & 0 & 4 \end{pmatrix}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 5 & 0 \end{pmatrix}\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} = \frac{27}{25}\epsilon. \tag{3.237}$$

The third method uses the probabilities:

$$E(0) = \sum_{n=1}^{2} P(E_n)E_n = \frac{8}{25}(-5)\epsilon + \frac{9}{25}(3)\epsilon + \frac{8}{25}(5)\epsilon = \frac{27}{25}\epsilon. \tag{3.238}$$

The energy at a time t is

$$\begin{aligned} E(t) = \langle\psi(t)|\hat{H}|\psi(t)\rangle &= \frac{8}{25}e^{iE_1t/\hbar}e^{-iE_1t/\hbar}\langle\phi_1|\hat{H}|\phi_1\rangle + \frac{9}{25}e^{iE_2t/\hbar}e^{-iE_2t/\hbar}\langle\phi_2|\hat{H}|\phi_2\rangle \\ &+ \frac{8}{25}e^{iE_3t/\hbar}e^{-iE_3t/\hbar}\langle\phi_3|\hat{H}|\phi_3\rangle = \frac{8}{25}(-5)\epsilon + \frac{9}{25}(3)\epsilon + \frac{8}{25}(5)\epsilon = \frac{27}{25}\epsilon = E(0). \end{aligned} \tag{3.239}$$

As expected, $E(t) = E(0)$ since $d\langle\hat{H}\rangle/dt = 0$.

(d) Since none of the eigenvalues of $\hat{H}$ is degenerate, the eigenvectors $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$ form a compete (orthonormal) basis. Thus $\{\hat{H}\}$ forms a complete set of commuting operators.

Problem 3.12

(a) Calculate the Poisson bracket between the x and y components of the classical orbital angular momentum.

(b) Calculate the commutator between the x and y components of the orbital angular momentum operator.

(c) Compare the results obtained in (a) and (b).

Solution

(a) Using the definition (3.121) we can write the Poisson bracket $\{l_x, l_y\}$ as

$$\{l_x, l_y\} = \sum_{j=1}^{3}\left(\frac{\partial l_x}{\partial q_j}\frac{\partial l_y}{\partial p_j} - \frac{\partial l_x}{\partial p_j}\frac{\partial l_y}{\partial q_j}\right), \tag{3.240}$$

where $q_1 = x$, $q_2 = y$, $q_3 = z$, $p_1 = p_x$, $p_2 = p_y$, and $p_3 = p_z$. Since $l_x = yp_z - zp_y$, $l_y = zp_x - xp_z$, $l_z = xp_y - yp_x$, the only partial derivatives that survive are $\partial l_x/\partial z = -p_y$, $\partial l_y/\partial p_z = -x$, $\partial l_x/\partial p_z = y$, and $\partial l_y/\partial z = p_x$. Thus, we have

$$\{l_x, l_y\} = \frac{\partial l_x}{\partial z}\frac{\partial l_y}{\partial p_z} - \frac{\partial l_x}{\partial p_z}\frac{\partial l_y}{\partial z} = xp_y - yp_x = l_z. \tag{3.241}$$

(b) The components of $\hat{\vec{L}}$ are listed in (3.31) to (3.33): $\hat{L}_x = \hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y$, $\hat{L}_y = \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z$, and $\hat{L}_Z = \hat{X}\hat{P}_y - \hat{Y}\hat{P}_x$. Since $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ mutually commute and so do $\hat{P}_x$, $\hat{P}_y$, and $\hat{P}_z$, we have

$$\begin{aligned}
[\hat{L}_x, \hat{L}_y] &= [\hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y,\ \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z] \\
&= [\hat{Y}\hat{P}_z, \hat{Z}\hat{P}_x] - [\hat{Y}\hat{P}_z, \hat{X}\hat{P}_z] - [\hat{Z}\hat{P}_y, \hat{Z}\hat{P}_x] + [\hat{Z}\hat{P}_y, \hat{X}\hat{P}_z] \\
&= \hat{Y}[\hat{P}_z, \hat{Z}]\hat{P}_x + \hat{X}[\hat{Z}, \hat{P}_z]\hat{P}_y = i\hbar(\hat{X}\hat{P}_y - \hat{Y}\hat{P}_x) \\
&= i\hbar\hat{L}_z.
\end{aligned} \tag{3.242}$$

(c) A comparison of (3.241) and (3.242) shows that

$$\{l_x, l_y\} = l_z \longrightarrow [\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z. \tag{3.243}$$

Problem 3.13

Consider a charged oscillator, of positive charge q and mass m, which is subject to an oscillating electric field $E_0 \cos \omega t$; the particle's Hamiltonian is $\hat{H} = P^2/(2m) + k\hat{X}^2/2 + qE_0\hat{X}\cos \omega t$.

(a) Calculate $d\langle\hat{X}\rangle/dt$, $d\langle\hat{P}\rangle/dt$, $d\langle\hat{H}\rangle/dt$.

(b) Solve the equation for $d\langle\hat{X}\rangle/dt$ and obtain $\langle\hat{X}\rangle(t)$ such that $\langle\hat{X}\rangle(0) = x_0$.

Solution

(a) Since the position operator $\hat{X}$ does not depend explicitly on time (i.e., $\partial\hat{X}/\partial t = 0$), equation (3.96) yields

$$\frac{d}{dt}\langle\hat{X}\rangle = \frac{1}{i\hbar}\langle[\hat{X},\hat{H}]\rangle = \frac{1}{i\hbar}\left\langle\left[\hat{X},\frac{\hat{P}^2}{2m}\right]\right\rangle = \frac{\langle\hat{P}\rangle}{m}. \tag{3.244}$$

Now, since $[\hat{P},\hat{X}] = -i\hbar$, $[\hat{P},\hat{X}^2] = -2i\hbar\hat{X}$ and $\partial\hat{P}/\partial t = 0$, we have

$$\frac{d}{dt}\langle\hat{P}\rangle = \frac{1}{i\hbar}\langle[\hat{P},\hat{H}]\rangle = \frac{1}{i\hbar}\left\langle\left[\hat{P},\frac{1}{2}k\hat{X}^2 + qE_0\hat{X}\cos\omega t\right]\right\rangle = -k\langle\hat{X}\rangle - qE_0\cos\omega t, \tag{3.245}$$

$$\frac{d}{dt}\langle\hat{H}\rangle = \frac{1}{i\hbar}\langle[\hat{H},\hat{H}]\rangle + \left\langle\frac{\partial\hat{H}}{\partial t}\right\rangle = \left\langle\frac{\partial\hat{H}}{\partial t}\right\rangle = -qE_0\omega\langle\hat{X}\rangle\sin\omega t. \tag{3.246}$$

(b) To find $\langle\hat{X}\rangle$ we need to take a time derivative of (3.244) and then make use of (3.245):

$$\frac{d^2}{dt^2}\langle\hat{X}\rangle = \frac{1}{m}\frac{d}{dt}\langle\hat{P}\rangle = -\frac{k}{m}\langle\hat{X}\rangle - \frac{qE_0}{m}\cos\omega t. \tag{3.247}$$

The solution of this equation is

$$\langle\hat{X}\rangle(t) = \langle\hat{X}\rangle(0)\cos\left(\sqrt{\frac{k}{m}}t\right) - \frac{qE_0}{m\omega}\sin\omega t + A, \tag{3.248}$$

where A is a constant which can be determined from the initial conditions; since $\langle\hat{X}\rangle(0) = x_0$ we have $A = 0$, and hence

$$\langle\hat{X}\rangle(t) = x_0\cos\left(\sqrt{\frac{k}{m}}t\right) - \frac{qE_0}{m\omega}\sin\omega t. \tag{3.249}$$

Problem 3.14

Consider a one-dimensional free particle of mass m whose position and momentum at time $t = 0$ are given by x_0 and p_0, respectively.

(a) Calculate $\langle\hat{P}\rangle(t)$ and show that $\langle\hat{X}\rangle(t) = p_0t^2/m + x_0$.

(b) Show that $d\langle\hat{X}^2\rangle/dt = 2\langle\hat{P}\hat{X}\rangle/m + i\hbar/m$ and $d\langle\hat{P}^2\rangle/dt = 0$.

(c) Show that the position and momentum fluctuations are related by $d^2(\Delta x)^2/dt^2 = 2(\Delta p)^2/m^2$ and that the solution to this equation is given by $(\Delta x)^2 = (\Delta p)_0^2t^2/m^2 + (\Delta x)_0^2$ where $(\Delta x)_0$ and $(\Delta p)_0$ are the initial fluctuations.

Solution

(a) From the Ehrenfest equations $d\langle\hat{P}\rangle/dt = \langle[\hat{P},\hat{V}(x,t)]\rangle/i\hbar$ as shown in (3.144), and since for a free particle $\hat{V}(x,t) = 0$, we see that $d\langle\hat{P}\rangle/dt = 0$. As expected this leads to $\langle\hat{P}\rangle(t) = p_0$, since the linear momentum of a free particle is conserved. Inserting $\langle\hat{P}\rangle = p_0$ into Ehrenfest's other equation $d\langle\hat{X}\rangle/dt = \langle\hat{P}\rangle/m$ (see (3.142)), we obtain

$$\frac{d\langle\hat{X}\rangle}{dt} = \frac{1}{m}p_0. \tag{3.250}$$

The solution of this equation with the initial condition $\langle\hat{X}\rangle(0) = x_0$ is

$$\langle\hat{X}\rangle(t) = \frac{p_0}{m}t + x_0. \tag{3.251}$$

(b) First, the proof of $d\langle\hat{P}^2\rangle/dt = 0$ is straightforward. Since $[\hat{P}^2, \hat{H}] = [\hat{P}^2, \hat{P}^2/2m] = 0$ and $\partial\hat{P}^2/\partial t = 0$ (the momentum operator does not depend on time), (3.133) yields

$$\frac{d}{dt}\langle\hat{P}^2\rangle = \frac{1}{i\hbar}\langle[\hat{P}^2, \hat{H}]\rangle + \langle\frac{\partial\hat{P}^2}{\partial t}\rangle = 0. \tag{3.252}$$

For $d\langle\hat{X}^2\rangle/dt$ we have

$$\frac{d}{dt}\langle\hat{X}^2\rangle = \frac{1}{i\hbar}\langle[\hat{X}^2, \hat{H}]\rangle = \frac{1}{2im\hbar}\langle[\hat{X}^2, \hat{P}^2]\rangle, \tag{3.253}$$

since $\partial\hat{X}^2/\partial t = 0$. Using $[\hat{X}, \hat{P}] = i\hbar$, we obtain

$$\begin{aligned}[\hat{X}^2, \hat{P}^2] &= \hat{P}[\hat{X}^2, \hat{P}] + [\hat{X}^2, \hat{P}]\hat{P} \\ &= \hat{P}\hat{X}[\hat{X}, \hat{P}] + \hat{P}[\hat{X}, \hat{P}]\hat{X} + \hat{X}[\hat{X}, \hat{P}]\hat{P} + [\hat{X}, \hat{P}]\hat{X}\hat{P} \\ &= 2i\hbar(\hat{P}\hat{X} + \hat{X}\hat{P}) = 2i\hbar(2\hat{P}\hat{X} + i\hbar);\end{aligned} \tag{3.254}$$

hence

$$\frac{d}{dt}\langle\hat{X}^2\rangle = \frac{2}{m}\langle\hat{P}\hat{X}\rangle + \frac{i\hbar}{m}. \tag{3.255}$$

(c) As the position fluctuation is given by $(\Delta x)^2 = \langle\hat{X}^2\rangle - \langle\hat{X}\rangle^2$, we have

$$\frac{d(\Delta x)^2}{dt} = \frac{d\langle\hat{X}^2\rangle}{dt} - 2\langle\hat{X}\rangle\frac{d\langle\hat{X}\rangle}{dt} = \frac{2}{m}\langle\hat{P}\hat{X}\rangle + \frac{i\hbar}{m} - \frac{2}{m}\langle\hat{X}\rangle\langle\hat{P}\rangle. \tag{3.256}$$

In deriving this expression we have used (3.255) and $d\langle\hat{X}\rangle/dt = \langle\hat{P}\rangle/m$. Now, since $d(\langle\hat{X}\rangle\langle\hat{P}\rangle)/dt = \langle\hat{P}\rangle d\langle\hat{X}\rangle/dt = \langle\hat{P}\rangle^2/m$ and

$$\frac{d\langle\hat{P}\hat{X}\rangle}{dt} = \frac{1}{i\hbar}\langle[\hat{P}\hat{X}, \hat{H}]\rangle = \frac{1}{2im\hbar}\langle[\hat{P}\hat{X}, \hat{P}^2]\rangle = \frac{1}{m}\langle\hat{P}^2\rangle, \tag{3.257}$$

we can write the second time derivative of (3.256) as follows:

$$\frac{d^2(\Delta x)^2}{dt^2} = \frac{2}{m}\left(\frac{d\langle\hat{P}\hat{X}\rangle}{dt} - \frac{d\langle\hat{X}\rangle\langle\hat{P}\rangle}{dt}\right) = \frac{2}{m^2}\left(\langle\hat{P}^2\rangle - \langle\hat{P}\rangle^2\right) = \frac{2}{m^2}(\Delta p)_0^2, \tag{3.258}$$

where $(\Delta p)_0^2 = \langle\hat{P}^2\rangle - \langle\hat{P}\rangle^2 = \langle\hat{P}^2\rangle_0 - \langle\hat{P}\rangle_0^2$; the momentum of the free particle is a constant of the motion. We can verify that the solution of the differential equation (3.258) is given by

$$(\Delta x)^2 = \frac{1}{m^2}(\Delta p)_0^2 t^2 + (\Delta x)_0^2. \tag{3.259}$$

This fluctuation is similar to the spreading of a Gaussian wave packet we derived in Chapter 1.

3.10 Exercises

Exercise 3.1
A particle in an infinite potential box with walls at $x = 0$ and $x = a$ (i.e., the potential is infinite for $x < 0$ and $x > a$ and zero in between) has the following wave function at some initial time:

$$\psi(x) = \frac{1}{\sqrt{5a}} \sin\left(\frac{\pi x}{a}\right) + \frac{2}{\sqrt{5a}} \sin\left(\frac{3\pi x}{a}\right).$$

(a) Find the possible results of the measurement of the system's energy and the corresponding probabilities.

(b) Find the form of the wave function after such a measurement.

(c) If the energy is measured again immediately afterwards, what are the relative probabilities of the possible outcomes?

Exercise 3.2
Let $\psi_n(x)$ denote the orthonormal stationary states of a system corresponding to the energy E_n. Suppose that the normalized wave function of the system at time $t = 0$ is $\psi(x, 0)$ and suppose that a measurement of the energy yields the value E_1 with probability 1/2, E_2 with probability 3/8, and E_3 with probability 1/8.

(a) Write the most general expansion for $\psi(x, 0)$ consistent with this information.

(b) What is the expansion for the wave function of the system at time t, $\psi(x, t)$?

(c) Show that the expectation value of the Hamiltonian does not change with time.

Exercise 3.3
Consider a neutron which is confined to an infinite potential well of width $a = 8$ fm. At time $t = 0$ the neutron is assumed to be in the state

$$\Psi(x, 0) = \sqrt{\frac{4}{7a}} \sin\left(\frac{\pi x}{a}\right) + \sqrt{\frac{2}{7a}} \sin\left(\frac{2\pi x}{a}\right) + \sqrt{\frac{8}{7a}} \sin\left(\frac{3\pi x}{a}\right).$$

(a) If an energy measurement is carried out on the system, what are the values that will be found for the energy and with what probabilities? Express your answer in MeV (the mass of the neutron is $mc^2 \simeq 939$ MeV, $\hbar c \simeq 197$ MeV fm).

(b) If this measurement is repeated on many identical systems, what is the average value of the energy that will be found? Again, express your answer in MeV.

(c) Using the uncertainty principle, estimate the order of magnitude of the neutron's speed in this well as a function of the speed of light c.

Exercise 3.4
Consider the dimensionless harmonic oscillator Hamiltonian

$$\hat{H} = \frac{1}{2}\hat{P}^2 + \frac{1}{2}\hat{X}^2, \quad \text{with } \hat{P} = -i\frac{d}{dx}.$$

(a) Show that the two wave functions $\psi_0(x) = e^{-x^2/2}$ and $\psi_1(x) = xe^{-x^2/2}$ are eigenfunctions of $\hat{H}$ with eigenvalues 1/2 and 3/2, respectively.

(b) Find the value of the coefficient α such that $\psi_2(x) = \left(1 + \alpha x^2\right)e^{-x^2/2}$ is orthogonal to $\psi_0(x)$. Then show that $\psi_2(x)$ is an eigenfunction of $\hat{H}$ with eigenvalue 5/2.

Exercise 3.5
Consider that the wave function of a dimensionless harmonic oscillator, whose Hamiltonian is $\hat{H} = \frac{1}{2}\hat{P}^2 + \frac{1}{2}\hat{X}^2$, is given at time $t = 0$ by

$$\psi(x,0) = \frac{1}{\sqrt{8\pi}}\phi_0(x) + \frac{1}{\sqrt{18\pi}}\phi_2(x) = \frac{1}{\sqrt{8\pi}}e^{-x^2/2} + \frac{1}{\sqrt{18\pi}}\left(1-2x^2\right)e^{-x^2/2}.$$

(a) Find the expression of the oscillator's wave function at any later time t.

(b) Calculate the probability P_0 to find the system in an eigenstate of energy 1/2 and the probability P_2 of finding the system in an eigenstate of energy 5/2.

(c) Calculate the probability density, $\rho(x,t)$, and the current density, $\vec{J}(x,t)$.

(d) Verify that the probability is conserved; that is, show that $\partial\rho/\partial t + \vec{\nabla}\cdot\vec{J}(x,t) = 0$.

Exercise 3.6
A particle of mass m, in an infinite potential well of length a, has the following initial wave function at $t = 0$:

$$\psi(x,0) = \sqrt{\frac{3}{5a}}\sin\left(\frac{3\pi x}{a}\right) + \frac{1}{\sqrt{5a}}\sin\left(\frac{5\pi x}{a}\right), \tag{3.260}$$

and an energy spectrum $E_n = -\hbar^2\pi^2n^2/(2ma^2)$.

Find $\psi(x,t)$ at any later time t, then calculate $\frac{\partial\rho}{\partial t}$ and the probability current density vector $\vec{J}(x,t)$ and verify that $\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{J}(x,t) = 0$. Recall that $\rho = \psi^*(x,t)\psi(x,t)$ and $\vec{J}(x,t) = \frac{i\hbar}{2m}\left(\psi(x,t)\vec{\nabla}\psi^*(x,t) - \psi^*(x,t)\vec{\nabla}\psi(x,t)\right)$.

Exercise 3.7
Consider a system whose initial state at $t = 0$ is given in terms of a complete and orthonormal set of three vectors: $|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle$ as follows: $|\psi(0)\rangle = 1/\sqrt{3}|\phi_1\rangle + A|\phi_2\rangle + 1/\sqrt{6}|\phi_3\rangle$, where A is a real constant.

(a) Find A so that $|\psi(0)\rangle$ is normalized.

(b) If the energies corresponding to $|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle$ are given by E_1, E_2, and E_3, respectively, write down the state of the system $|\psi(t)\rangle$ at any later time t.

(c) Determine the probability of finding the system at a time t in the state $|\phi_3\rangle$.

Exercise 3.8
The components of the initial state $|\psi_i\rangle$ of a quantum system are given in a complete and orthonormal basis of three states $|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle$ by

$$\langle\phi_1|\psi_i\rangle = \frac{i}{\sqrt{3}}, \quad \langle\phi_2|\psi_i\rangle = \sqrt{\frac{2}{3}}, \quad \langle\phi_3|\psi_i\rangle = 0.$$

Calculate the probability of finding the system in a state $|\psi_f\rangle$ whose components are given in the same basis by

$$\langle\phi_1|\psi_f\rangle = \frac{1+i}{\sqrt{3}}, \quad \langle\phi_2|\psi_f\rangle = \frac{1}{\sqrt{6}}, \quad \langle\phi_3|\psi_f\rangle = \frac{1}{\sqrt{6}}.$$

Exercise 3.9

(a) Evaluate the Poisson bracket $\{x^2, p^2\}$.

(b) Express the commutator $\left[\hat{X}^2, \hat{P}^2\right]$ in terms of $\hat{X}\hat{P}$ plus a constant in $\hbar^2$.

(c) Find the classical limit of $\left[\hat{x}^2, \hat{p}^2\right]$ for this expression and then compare it with the result of part (a).

Exercise 3.10

A particle bound in a one-dimensional potential has a wave function

$$\psi(x) = \begin{cases} Ae^{5ikx}\cos(3\pi x/a), & -a/2 \le x \le a/2, \\ 0, & |x| > a/2. \end{cases}$$

(a) Calculate the constant A so that $\psi(x)$ is normalized.

(b) Calculate the probability of finding the particle between $x = 0$ and $x = a/4$.

Exercise 3.11

(a) Show that any component of the momentum operator of a particle is compatible with its kinetic energy operator.

(b) Show that the momentum operator is compatible with the Hamiltonian operator only if the potential operator is constant in space coordinates.

Exercise 3.12

Consider a physical system whose Hamiltonian H and an operator A are given by

$$H = \mathcal{E}_0\begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A = a_0\begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix},$$

where $\mathcal{E}_0$ has the dimensions of energy.

(a) Do H and A commute? If yes, give a basis of eigenvectors common to H and A.

(b) Which among the sets of operators $\{\hat{H}\}$, $\{\hat{A}\}$, $\{\hat{H}, \hat{A}\}$, $\{\hat{H}^2, \hat{A}\}$ form a complete set of commuting operators (CSCO)?

Exercise 3.13

Show that the momentum and the total energy can be measured simultaneously only when the potential is constant everywhere.

Exercise 3.14

The initial state of a system is given in terms of four orthonormal energy eigenfunctions $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$, and $|\phi_4\rangle$ as follows:

$$|\psi_0\rangle = |\psi(t=0)\rangle = \frac{1}{\sqrt{3}}|\phi_1\rangle + \frac{1}{2}|\phi_2\rangle + \frac{1}{\sqrt{6}}|\phi_3\rangle + \frac{1}{2}|\phi_4\rangle.$$

(a) If the four kets $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$, and $|\phi_4\rangle$ are eigenvectors to the Hamiltonian $\hat{H}$ with energies E_1, E_2, E_3, and E_4, respectively, find the state $|\psi(t)\rangle$ at any later time t.

(b) What are the possible results of measuring the energy of this system and with what probability will they occur?

(c) Find the expectation value of the system's Hamiltonian at $t = 0$ and $t = 10\,\text{s}$.

Exercise 3.15
The complete set expansion of an initial wave function $\psi(x,0)$ of a system in terms of orthonormal energy eigenfunctions $\phi_n(x)$ of the system has three terms, $n = 1,2,3$. The measurement of energy on the system represented by $\psi(x,0)$ gives three values, E_1 and E_2 with probability 1/4 and E_3 with probability 1/2.

(a) Write down $\psi(x,0)$ in terms of $\phi_1(x)$, $\phi_2(x)$, and $\phi_3(x)$.

(b) Find $\psi(x,0)$ at any later time t, i.e., find $\psi(x,t)$.

Exercise 3.16
Consider a system whose Hamiltonian H and an operator A are given by the matrices

$$H = \mathcal{E}_0 \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 2i \\ 0 & -2i & 0 \end{pmatrix}, \quad A = a_0 \begin{pmatrix} 0 & -i & 0 \\ i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

(a) If we measure energy, what values will we obtain?

(b) Suppose that when we measure energy, we obtain a value of $\sqrt{5}\mathcal{E}_0$. Immediately afterwards, we measure A. What values will we obtain for A and what are the probabilities corresponding to each value?

(c) Calculate the expectation value $\langle A \rangle$.

Exercise 3.17
Consider a physical system whose Hamiltonian and initial state are given by

$$H = \mathcal{E}_0 \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad |\psi_0\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix},$$

where $\mathcal{E}_0$ has the dimensions of energy.

(a) What values will we obtain when measuring the energy and with what probabilities?

(b) Calculate the expectation value of the Hamiltonian $\langle \hat{H} \rangle$.

Exercise 3.18
Consider a system whose state $|\psi(t)\rangle$ and two observables A and B are given by

$$|\psi(t)\rangle = \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}, \qquad A = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

(a) We perform a measurement where A is measured first and then B immediately afterwards. Find the probability of obtaining a value of $\sqrt{2}$ for A and a value of -1 for B.

(b) Now we measure B first and then A immediately afterwards. Find the probability of obtaining a value of -1 for B and a value of $\sqrt{2}$ for A.

(c) Compare the results of (a) and (b). Explain.

Exercise 3.19
Consider a system whose state $|\psi(t)\rangle$ and two observables A and B are given by

$$|\psi(t)\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} -i \\ 2 \\ 0 \end{pmatrix}, \qquad A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i & 1 \\ -i & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & i \\ 0 & -i & 0 \end{pmatrix}.$$

(a) Are A and B compatible? Which among the sets of operators $\{\hat{A}\}$, $\{\hat{B}\}$, and $\{\hat{A}, \hat{B}\}$ form a complete set of commuting operators?

(b) Measuring A first and then B immediately afterwards, find the probability of obtaining a value of -1 for A and a value of 3 for B.

(c) Now, measuring B first then A immediately afterwards, find the probability of obtaining 3 for B and -1 for A. Compare this result with the probability obtained in (b).

Exercise 3.20

Consider a physical system which has a number of observables that are represented by the following matrices:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & i \\ -1 & -i & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 3 & 1 \end{pmatrix}.$$

(a) Find the results of the measurements of the compatible observables.

(b) Which among these observables are compatible? Give a basis of eigenvectors common to these observables.

(c) Which among the sets of operators $\{\hat{A}\}$, $\{\hat{B}\}$, $\{\hat{C}\}$, $\{\hat{A}, \hat{B}\}$, $\{\hat{A}, \hat{C}\}$, $\{\hat{B}, \hat{C}\}$ form a complete set of commuting operators?

Exercise 3.21

Consider a system which is initially in a state $|\psi(0)\rangle$ and having a Hamiltonian $\hat{H}$, where

$$|\psi(0)\rangle = \begin{pmatrix} 4-i \\ -2+5i \\ 3+2i \end{pmatrix}, \qquad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 3 & 3 \\ 0 & 3 & 0 \end{pmatrix}.$$

(a) If a measurement of H is carried out, what values will we obtain and with what probabilities?

(b) Find the state of the system at a later time t; you may need to expand $|\psi(0)\rangle$ in terms of the eigenvectors of H.

(c) Find the total energy of the system at time $t = 0$ and any later time t; are these values different?

(d) Does $\hat{H}$ form a complete set of commuting operators?

Exercise 3.22

Consider a particle which moves in a scalar potential $V(\vec{r}) = V_x(x) + V_y(y) + V_z(z)$.

(a) Show that the Hamiltonian of this particle can be written as $\hat{H} = \hat{H}_x + \hat{H}_y + \hat{H}_z$, where $\hat{H}_x = p_x^2/(2m) + V_x(x)$, and so on.

(b) Do $\hat{H}_x$, $\hat{H}_y$, and $\hat{H}_z$ form a complete set of commuting operators?

Exercise 3.23

Consider a system whose Hamiltonian is $H = \mathcal{E}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, where $\mathcal{E}$ is a real constant with the dimensions of energy.

(a) Find the eigenenergies, E_1 and E_2, of H.

(b) If the system is initially (i.e., $t = 0$) in the state $|\psi_0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, find the probability so that a measurement of energy at $t = 0$ yields: (i) E_1, and (ii) E_2.

(c) Find the average value of the energy $\langle \hat{H} \rangle$ and the energy uncertainty $\sqrt{\langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2}$.

(d) Find the state $|\psi(t)\rangle$.

Exercise 3.24

Prove the relation

$$\frac{d}{dt}\langle \hat{A}\hat{B} \rangle = \langle \frac{\partial \hat{A}}{\partial t}\hat{B} \rangle + \langle \hat{A}\frac{\partial \hat{B}}{\partial t} \rangle + \frac{1}{i\hbar}\langle [\hat{A}, \hat{H}]\hat{B} \rangle + \frac{1}{i\hbar}\langle \hat{A}[\hat{B}, \hat{H}] \rangle.$$

Exercise 3.25

Consider a particle of mass m which moves under the influence of gravity; the particle's Hamiltonian is $\hat{H} = \hat{P}_z^2/(2m) - mg\hat{Z}$, where g is the acceleration due to gravity, $g = 9.8\,\mathrm{m\,s^{-2}}$.

(a) Calculate $d\langle \hat{Z} \rangle/dt$, $d\langle \hat{P}_z \rangle/dt$, $d\langle \hat{H} \rangle/dt$.

(b) Solve the equation $d\langle \hat{Z} \rangle/dt$ and obtain $\langle \hat{Z} \rangle(t)$, such that $\langle \hat{Z} \rangle(0) = h$ and $\langle \hat{P}_z \rangle(0) = 0$. Compare the result with the classical relation $z(t) = -\frac{1}{2}gt^2 + h$.

Exercise 3.26

Calculate $d\langle \hat{X} \rangle/dt$, $d\langle \hat{P}_x \rangle/dt$, $d\langle \hat{H} \rangle/dt$ for a particle with $\hat{H} = \hat{P}_x^2/(2m) + \frac{1}{2}m\omega^2\hat{X}^2 + V_0\hat{X}^3$.

Exercise 3.27

Consider a system whose initial state at $t = 0$ is given in terms of a complete and orthonormal set of four vectors $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$, $|\phi_4\rangle$ as follows:

$$|\psi(0)\rangle = \frac{A}{\sqrt{12}}|\phi_1\rangle + \frac{1}{\sqrt{6}}|\phi_2\rangle + \frac{2}{\sqrt{12}}|\phi_3\rangle + \frac{1}{2}|\phi_4\rangle,$$

where A is a real constant.

(a) Find A so that $|\psi(0)\rangle$ is normalized.

(b) If the energies corresponding to $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$, $|\phi_4\rangle$ are given by E_1, E_2, E_3, and E_4, respectively, write down the state of the system $|\psi(t)\rangle$ at any later time t.

(c) Determine the probability of finding the system at a time t in the state $|\phi_2\rangle$.

Chapter 4

One-Dimensional Problems

4.1 Introduction

After presenting the formalism of quantum mechanics in the previous two chapters, we are now well equipped to apply it to the study of physical problems. Here we apply the Schrödinger equation to *one-dimensional* problems. These problems are interesting since there exist many physical phenomena whose motion is one-dimensional. The application of the Schrödinger equation to *one-dimensional* problems enables us to compare the predictions of classical and quantum mechanics in a simple setting. In addition to being simple to solve, one-dimensional problems will be used to illustrate some nonclassical effects.

The Schrödinger equation describing the dynamics of a microscopic particle of mass m in a one-dimensional time-independent potential $V(x)$ is given by

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x), \tag{4.1}$$

where E is the total energy of the particle. The solutions of this equation yield the allowed energy eigenvalues E_n and the corresponding wave functions $\psi_n(x)$. To solve this partial differential equation, we need to specify the potential $V(x)$ as well as the boundary conditions; the boundary conditions can be obtained from the physical requirements of the system.

We have seen in the previous chapter that the solutions of the Schrödinger equation for time-independent potentials are stationary,

$$\Psi(x,t) = \psi(x)e^{-iEt/\hbar}, \tag{4.2}$$

for the probability density does not depend on time. Recall that the state $\psi(x)$ has the physical dimensions of $1/\sqrt{L}$, where L is a length. Hence, the physical dimension of $|\psi(x)|^2$ is $1/L$: $\left[|\psi(x)|^2\right] = 1/L$.

We begin by examining some general properties of one-dimensional motion and discussing the symmetry character of the solutions. Then, in the rest of the chapter, we apply the Schrödinger equation to various one-dimensional potentials: the free particle, the potential step, the finite and infinite potential wells, and the harmonic oscillator. We conclude by showing how to solve the Schrödinger equation numerically.

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

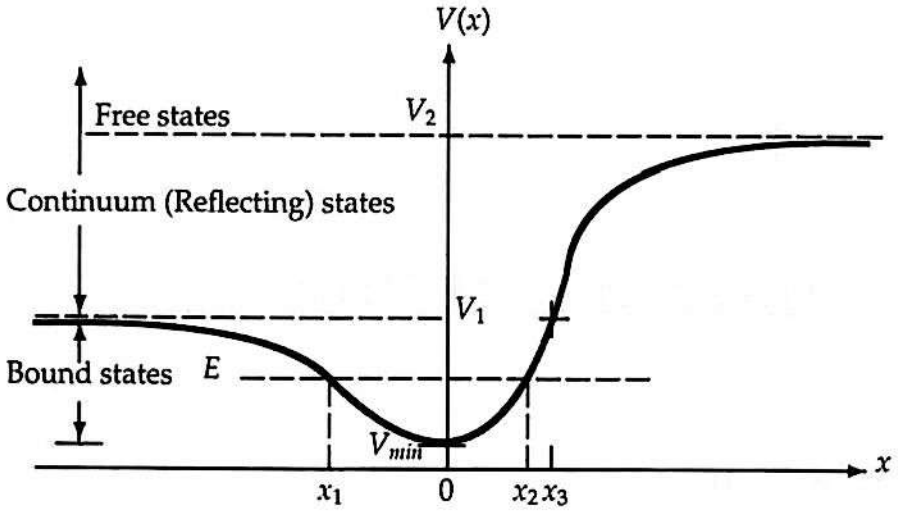


Figure 4.1: Shape of a general potential displaying discrete (bound) and continuous (reflecting and free) states.

4.2 Properties of One-Dimensional Motion

To study the dynamic properties of a single particle moving in a one-dimensional potential, let us consider a potential $V(x)$ that is general enough to allow for the illustration of all the desired features. One such potential is displayed in Figure 4.1; it is finite at $x \to \pm\infty$, $V(-\infty) = V_1$ and $V(+\infty) = V_2$ with V_1 smaller than V_2, and it has a minimum, V_{min}. In particular, we want to study the conditions under which discrete and continuous spectra occur. As the character of the states is completely determined by the size of the system's energy, we will be considering separately the cases where the energy is smaller and larger than the potential.

4.2.1 Discrete Spectrum (Bound States)

Bound states occur whenever the particle cannot move to infinity. That is, the particle is *confined* or *bound* at all energies to move within a *finite* and limited region of space which is delimited by two classical turning points. The Schrödinger equation in this region admits only solutions that are *discrete*. The infinite square well potential and the harmonic oscillator are typical examples that display bound states.

In the potential of Figure 4.1, the motion of the particle is bounded between the classical turning points x_1 and x_2 when the particle's energy lies between V_{min} and V_1:

$$V_{min} < E < V_1. \tag{4.3}$$

The states corresponding to this energy range are called *bound* states. They are defined as states whose wave functions are finite (or zero) at $x \to \pm\infty$; usually the bound states have energies smaller than the potential $E < V$. For the bound states to exist, the potential $V(x)$ must have at least one minimum which is lower than V_1 (i.e., $V_{min} < V_1$). The energy spectra of bound states are discrete. We need to use the boundary conditions[1] to find the wave function and the energy.

[1]Since the Schrödinger equation is a *second-order* differential equation, only *two* boundary conditions are required to solve it.

Let us now list two theorems that are important to the study of bound states.

Theorem 4.1 *In a one-dimensional problem the energy levels of a bound state system are discrete and not degenerate.*

Theorem 4.2 *The wave function $\psi_n(x)$ of a one-dimensional bound state system has n nodes (i.e., $\psi_n(x)$ vanishes n times) if $n = 0$ corresponds to the ground state and $(n-1)$ nodes if $n = 1$ corresponds to the ground state.*

4.2.2 Continuous Spectrum (Unbound States)

Unbound states occur in those cases where the motion of the system is not confined; a typical example is the *free* particle. For the potential displayed in Figure 4.1 there are two energy ranges where the particle's motion is infinite: $V_1 < E < V_2$ and $E > V_2$.

- Case $V_1 < E < V_2$

  In this case the particle's motion is infinite only towards $x = -\infty$. That is, the particle coming from $x = -\infty$ gets reflected at $x = x_3$ (the classical turning point) back to $x \to -\infty$; we will have an oscillatory motion between $-\infty < x < x_3$. The energy spectrum is continuous and none of the energy eigenvalues is degenerate. The nondegeneracy can be shown to result as follows. Since the Schrödinger equation (4.1) is a second-order differential equation, it has, for this case, two linearly independent solutions, but only one is physically acceptable. The solution is oscillatory for $x \leq x_3$ and rapidly decaying for $x > x_3$ so that it is finite (zero) at $x \to +\infty$, since divergent solutions are unphysical.

- Case $E > V_2$

  The energy spectrum is continuous and the particle's motion is infinite in both directions of x (i.e., towards $x \to \pm\infty$). All the energy levels of this spectrum are doubly degenerate. To see this, note that the general solution to (4.1) is a linear combination of two independent oscillatory solutions, one moving to the left and the other to the right. In the previous nondegenerate case only one solution is retained, since the other one diverges as $x \to +\infty$ and it has to be rejected.

In contrast to bound states, unbound states cannot be normalized and we cannot use boundary conditions.

4.2.3 Mixed Spectrum

Potentials that confine the particle for only some energies give rise to mixed spectra; the motion of the particle for such potentials is confined for some energy values only. For instance, for the potential displayed in Figure 4.1, if the energy of the particle is between $V_{min} < E < V_1$, the motion of the particle is confined (bound) and its spectrum is discrete, but if $E > V_2$, the particle's motion is unbound and its spectrum is continuous (if $V_1 < E < V_2$, the motion is unbound only along the $x = -\infty$ direction). Other typical examples where mixed spectra are encountered are the finite square well potential and the Coulomb or molecular potential.

4.2.4 Symmetric Potentials and Parity

Most of the potentials that are encountered at the microscopic level are symmetric (or even) with respect to space inversion, $\hat{V}(-x) = \hat{V}(x)$. This symmetry introduces considerable simplifications in the calculations.

When $\hat{V}(x)$ is even, the corresponding Hamiltonian, $\hat{H}(x) = -(\hbar^2/2m)d^2/dx^2 + \hat{V}(x)$, is also even. We saw in Chapter 2 that even operators commute with the parity operator; hence they can have a common eigenbasis.

Let us consider the following two cases pertaining to degenerate and nondegenerate spectra of this Hamiltonian:

- **Nondegenerate spectrum**

  First we consider the particular case where the eigenvalues of the Hamiltonian corresponding to this symmetric potential are not degenerate. According to Theorem 4.1, this Hamiltonian describes bound states. We saw in Chapter 2 that a nondegenerate, even operator has the same eigenstates as the parity operator. Since the eigenstates of the parity operator have definite parity, *the bound eigenstates of a particle moving in a one-dimensional symmetric potential have definite parity; they are either even or odd*:

$$\boxed{\hat{V}(-x) = \hat{V}(x) \implies \psi(-x) = \pm\psi(x).} \tag{4.4}$$

- **Degenerate spectrum**

  If the spectrum of the Hamiltonian corresponding to a symmetric potential is degenerate, the eigenstates are expressed only in terms of even and odd states. That is, the eigenstates do not have definite parity.

Summary: The various properties of the one-dimensional motion discussed in this section can be summarized as follows:

- The energy spectrum of a bound state system is discrete and nondegenerate.
- The bound state wave function $\psi_n(x)$ has: (a) n nodes if $n = 0$ corresponds to the ground state and (b) $(n-1)$ nodes if $n = 1$ corresponds to the ground state.
- The bound state eigenfunctions in an even potential have definite parity.
- The eigenfunctions of a degenerate spectrum in an even potential do not have definite parity.

4.3 The Free Particle: Continuous States

This is the simplest one-dimensional problem; it corresponds to $V(x) = 0$ for any value of x. In this case the Schrödinger equation is given by

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} = E\psi(x) \qquad \Longrightarrow \qquad \left(\frac{d^2}{dx^2} + k^2\right)\psi(x) = 0, \tag{4.5}$$

where $k^2 = 2mE/\hbar^2$, k being the wave number. The most general solution to (4.5) is a combination of two linearly independent *plane waves* $\psi_+(x) = e^{ikx}$ and $\psi_-(x) = e^{-ikx}$:

$$\psi_k(x) = A_+e^{ikx} + A_-e^{-ikx}, \tag{4.6}$$

where A_+ and A_- are two arbitrary constants. The complete wave function is thus given by the stationary state

$$\Psi_k(x,t) = A_+e^{i(kx-\omega t)} + A_-e^{-i(kx+\omega t)} = A_+e^{i(kx-\hbar k^2t/2m)} + A_-e^{-i(kx+\hbar k^2t/2m)}, \tag{4.7}$$

since $\omega = E/\hbar = \hbar k^2/2m$. The first term, $\Psi_+(x,t) = A_+e^{i(kx-\omega t)}$, represents a wave traveling to the right, while the second term, $\Psi_-(x,t) = A_-e^{-i(kx+\omega t)}$, represents a wave traveling to the left. The intensities of these waves are given by $|A_+|^2$ and $|A_-|^2$, respectively. We should note that the waves $\Psi_+(x,t)$ and $\Psi_-(x,t)$ are associated, respectively, with a free particle traveling to the right and to the left with *well-defined* momenta and energy: $p_\pm = \pm\hbar k$, $E_\pm = \hbar^2k^2/2m$. We will comment on the physical implications of this in a moment. Since there are no boundary conditions, there are no restrictions on k or on E; all values yield solutions to the equation.

The free particle problem is simple to solve mathematically, yet it presents a number of physical subtleties. Let us discuss briefly three of these subtleties. First, the probability densities corresponding to either solutions

$$P_\pm(x,t) = |\Psi_\pm(x,t)|^2 = |A_\pm|^2 \tag{4.8}$$

are constant, for they depend neither on x nor on t. This is due to the complete loss of information about the position and time for a state with definite values of momentum, $p_\pm = \pm\hbar k$, and energy, $E_\pm = \hbar^2k^2/2m$. This is a consequence of Heisenberg's uncertainty principle: when the momentum and energy of a particle are known exactly, $\Delta p = 0$ and $\Delta E = 0$, there must be total uncertainty about its position and time: $\Delta x \longrightarrow \infty$ and $\Delta t \longrightarrow \infty$. The second subtlety pertains to an apparent discrepancy between the speed of the wave and the speed of the particle it is supposed to represent. The speed of the plane waves $\Psi_\pm(x,t)$ is given by

$$v_{wave} = \frac{\omega}{k} = \frac{E}{\hbar k} = \frac{\hbar^2k^2/2m}{\hbar k} = \frac{\hbar k}{2m}. \tag{4.9}$$

On the other hand, the classical speed of the particle[2] is given by

$$v_{classical} = \frac{p}{m} = \frac{\hbar k}{m} = 2v_{wave}. \tag{4.10}$$

This means that the particle travels twice as fast as the wave that represents it! Third, the wave function is not normalizable:

$$\int_{-\infty}^{+\infty} \Psi_\pm^*(x,t)\Psi_\pm(x,t)\,dx = |A_\pm|^2 \int_{-\infty}^{+\infty} dx \to \infty. \tag{4.11}$$

The solutions $\Psi_\pm(x,t)$ are thus unphysical; physical wave functions must be square integrable. The problem can be traced to this: a free particle cannot have sharply defined momenta and energy.

[2]The classical speed can be associated with the flux (or current density) which, as shown in Chapter 3, is $J_+ = i\hbar\frac{1}{2m}(\Psi_+\frac{\partial\Psi_+^*}{\partial x} - \Psi_+^*\frac{\partial\Psi_+}{\partial x}) = \frac{\hbar k}{m} = \frac{p}{m}$, where use was made of $A_+ = 1$.

In view of the three subtleties outlined above, the solutions of the Schrödinger equation (4.5) that are physically acceptable cannot be plane waves. Instead, we can construct physical solutions by means of a linear superposition of plane waves. The answer is provided by *wave packets*, which we have seen in Chapter 1:

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{i(kx-\omega t)} dk, \tag{4.12}$$

where $\phi(k)$, the amplitude of the wave packet, is given by the Fourier transform of $\psi(x,0)$ as

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \psi(x,0) e^{-ikx} dx. \tag{4.13}$$

The wave packet solution cures and avoids all the subtleties raised above. First, the momentum, the position and the energy of the particle are no longer known exactly; only probabilistic outcomes are possible. Second, as shown in Chapter 1, the wave packet (4.12) and the particle travel with the same speed $v_g = p/m$, called the *group* speed or the speed of the whole packet. Third, the wave packet (4.12) is normalizable.

To summarize, a free particle cannot be represented by a single (monochromatic) plane wave; it has to be represented by a wave packet. The physical solutions of the Schrödinger equation are thus given by wave packets, not by stationary solutions.

4.4 The Potential Step

Another simple problem consists of a particle that is free everywhere, but beyond a particular point, say $x = 0$, the potential increases sharply (i.e., it becomes repulsive or attractive). A potential of this type is called a potential step (see Figure 4.2):

$$V(x) = \begin{cases} 0, & x < 0, \\ V_0, & x \geq 0. \end{cases} \tag{4.14}$$

In this problem we try to analyze the dynamics of a flux of particles (all having the same mass m and moving with the same velocity) moving from left to the right. We are going to consider two cases, depending on whether the energy of the particles is larger or smaller than V_0.

(a) Case $E > V_0$

The particles are free for $x < 0$ and feel a repulsive potential V_0 that starts at $x = 0$ and stays flat (constant) for $x > 0$. Let us analyze the dynamics of this flux of particles classically and then quantum mechanically.

Classically, the particles approach the potential step or barrier from the left with a constant momentum $\sqrt{2mE}$. As the particles enter the region $x \geq 0$, where the potential now is $V = V_0$, they slow down to a momentum $\sqrt{2m(E - V_0)}$; they will then conserve this momentum as they travel to the right. Since the particles have sufficient energy to penetrate into the region $x \geq 0$, there will be *total transmission*: all the particles will emerge to the right with a smaller kinetic energy $E - V_0$. This is then a simple *scattering* problem in one dimension.

Quantum mechanically, the dynamics of the particle is regulated by the Schrödinger equation, which is given in these two regions by

$$\left(\frac{d^2}{dx^2} + k_1^2\right) \psi_1(x) = 0 \qquad (x < 0), \tag{4.15}$$

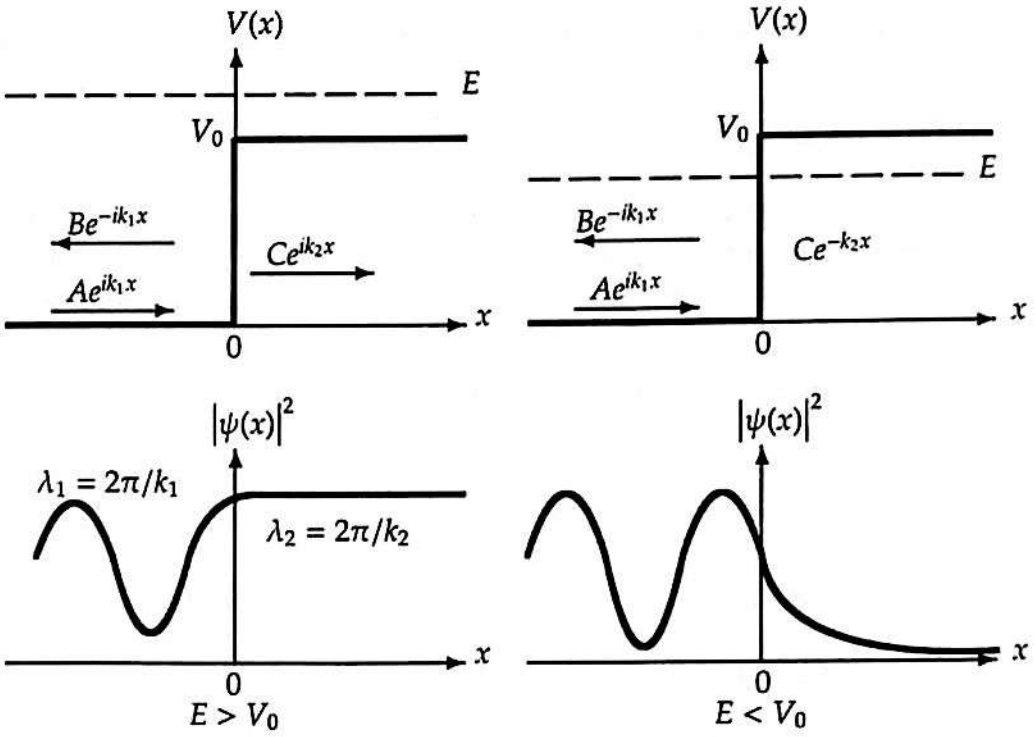


Figure 4.2: Potential step and propagation directions of the incident, reflected, and transmitted waves, plus their probability densities $|\psi(x)|^2$ when $E > V_0$ and $E < V_0$.

$$\left(\frac{d^2}{dx^2} + k_2^2\right)\psi_2(x) = 0 \qquad (x \geq 0), \tag{4.16}$$

where $k_1^2 = 2mE/\hbar^2$ and $k_2^2 = 2m(E - V_0)/\hbar^2$. The most general solutions to these two equations are plane waves:

$$\psi_1(x) = Ae^{ik_1x} + Be^{-ik_1x} \qquad (x < 0), \tag{4.17}$$

$$\psi_2(x) = Ce^{ik_2x} + De^{-ik_2x} \qquad (x \geq 0), \tag{4.18}$$

where Ae^{ik_1x} and Ce^{ik_2x} represent waves moving in the positive x-direction, but Be^{-ik_1x} and De^{-ik_2x} correspond to waves moving in the negative x-direction. We are interested in the case where the particles are initially incident on the potential step from the left: they can be reflected or transmitted at $x = 0$. Since no wave is reflected from the region $x > 0$ to the left, the constant D must vanish. Since we are dealing with stationary states, the complete wave function is thus given by

$$\Psi(x,t) = \begin{cases} \psi_1(x)e^{-i\omega t} = Ae^{i(k_1x-\omega t)} + Be^{-i(k_1x+\omega t)} & x < 0 \\ \psi_2(x)e^{-i\omega t} = Ce^{i(k_2x-\omega t)} & x \geq 0, \end{cases} \tag{4.19}$$

where $A\exp[i(k_1x - \omega t)]$, $B\exp[-i(k_1x + \omega t)]$, and $C\exp[i(k_2x - \omega t)]$ represent the *incident*, the *reflected*, and the *transmitted* waves, respectively; they travel to the right, the left, and the right (Figure 4.2). Note that the probability density $|\psi(x)|^2$ shown in the lower left plot of Figure 4.2 is a straight line for $x > 0$, since $|\psi_2(x)|^2 = |C\exp i(k_2x - \omega t)|^2 = |C|^2$.

Let us now evaluate the *reflection* and *transmission coefficients*, R and T, as defined by

$$R = \left|\frac{\text{reflected current density}}{\text{incident current density}}\right| = \left|\frac{J_{reflected}}{J_{incident}}\right|, \qquad T = \left|\frac{J_{transmitted}}{J_{incident}}\right|; \tag{4.20}$$

R represents the ratio of the reflected to the incident beams and T the ratio of the transmitted to the incident beams. To calculate R and T, we need to find $J_{incident}$, $J_{reflected}$, and $J_{transmitted}$. Since the incident wave is $\psi_i(x) = Ae^{ik_1x}$, the incident current density (or incident flux) is given by

$$J_{incident} = \frac{i\hbar}{2m}\left(\psi_i(x)\frac{d\psi_i^*(x)}{dx} - \psi_i^*(x)\frac{d\psi_i(x)}{dx}\right) = \frac{\hbar k_1}{m}|A|^2. \tag{4.21}$$

Similarly, since the reflected and transmitted waves are $\psi_r(x) = Be^{-ik_1x}$ and $\psi_t(x) = Ce^{ik_2x}$, we can verify that the reflected and transmitted fluxes are

$$J_{reflected} = -\frac{\hbar k_1}{m}|B|^2, \qquad J_{transmitted} = \frac{\hbar k_2}{m}|C|^2. \tag{4.22}$$

A combination of (4.20) to (4.22) yields

$$R = \frac{|B|^2}{|A|^2}, \qquad T = \frac{k_2}{k_1}\frac{|C|^2}{|A|^2}. \tag{4.23}$$

Thus, the calculation of R and T is reduced to determining the constants B and C. For this, we need to use the boundary conditions of the wave function at $x = 0$. Since both the wave function and its first derivative are continuous at $x = 0$,

$$\psi_1(0) = \psi_2(0), \qquad \frac{d\psi_1(0)}{dx} = \frac{d\psi_2(0)}{dx}, \tag{4.24}$$

equations (4.17) and (4.18) yield

$$A + B = C, \qquad k_1(A - B) = k_2C; \tag{4.25}$$

hence

$$B = \frac{k_1 - k_2}{k_1 + k_2}A, \qquad C = \frac{2k_1}{k_1 + k_2}A. \tag{4.26}$$

As for the constant A, it can be determined from the normalization condition of the wave function, but we don't need it here, since R and T are expressed in terms of ratios. A combination of (4.23) with (4.26) leads to

$$R = \frac{(k_1 - k_2)^2}{(k_1 + k_2)^2} = \frac{(1 - \mathcal{K})^2}{(1 + \mathcal{K})^2}, \qquad T = \frac{4k_1k_2}{(k_1 + k_2)^2} = \frac{4\mathcal{K}}{(1 + \mathcal{K})^2}, \tag{4.27}$$

where $\mathcal{K} = k_2/k_1 = \sqrt{1 - V_0/E}$. The sum of R and T is equal to 1, as it should be.

In contrast to classical mechanics, which states that none of the particles get reflected, equation (4.27) shows that the quantum mechanical reflection coefficient R is not zero: there are particles that get reflected in spite of their energies being higher than the step V_0. This effect must be attributed to the *wavelike behavior* of the particles.

From (4.27) we see that as E gets smaller and smaller, T also gets smaller and smaller so that when $E = V_0$ the transmission coefficient T becomes zero and $R = 1$. On the other hand, when $E \gg V_0$, we have $\mathcal{K} = \sqrt{1 - V_0/E} \simeq 1$; hence $R = 0$ and $T = 1$. This is expected since, when the incident particles have very high energies, the potential step is so weak that it produces no noticeable effect on their motion.

Remark: physical meaning of the boundary conditions

Throughout this chapter, we will encounter at numerous times the use of the boundary conditions of the wave function and its first derivative as in Eq (4.24). What is the underlying physics behind these continuity conditions? We can make two observations:

- Since the probability density $|\psi(x)|^2$ of finding the particle in any small region varies continuously from one point to another, the wave function $\psi(x)$ must, therefore, be a continuous function of x; thus, as shown in (4.24), we must have $\psi_1(0) = \psi_2(0)$.
- Since the linear momentum of the particle, $\hat{P}\psi(x) = -i\hbar d\psi(x)/dx$, must be a continuous function of x as the particle moves from left to right, the first derivative of the wave function, $d\psi(x)/dx$, must also be a continuous function of x, notably at $x = 0$. Hence, as shown in (4.24), we must have $d\psi_1(0)/dx = d\psi_2(0)/dx$.

(b) Case $E < V_0$

Classically, the particles arriving at the potential step from the left (with momenta $p = \sqrt{2mE}$) will come to a stop at $x = 0$ and then all will bounce back to the left with the magnitudes of their momenta unchanged. None of the particles will make it into the right side of the barrier $x = 0$; there is total reflection of the particles. So the motion of the particles is reversed by the potential barrier.

Quantum mechanically, the picture will be somewhat different. In this case, the Schrödinger equation and the wave function in the region $x < 0$ are given by (4.15) and (4.17), respectively. But for $x > 0$ the Schrödinger equation is given by

$$\left(\frac{d^2}{dx^2} - k_2'^2\right)\psi_2(x) = 0 \quad (x \geq 0), \tag{4.28}$$

where $k_2'^2 = 2m(V_0 - E)/\hbar^2$. This equation's solution is

$$\psi_2(x) = Ce^{-k_2'x} + De^{k_2'x} \quad (x \geq 0). \tag{4.29}$$

Since the wave function must be finite everywhere, and since the term $e^{k_2'x}$ diverges when $x \to \infty$, the constant D has to be zero. Thus, the complete wave function is

$$\Psi(x,t) = \begin{cases} Ae^{i(k_1x-\omega t)} + Be^{-i(k_1x+\omega t)}, & x < 0, \\ Ce^{-k_2'x}e^{-i\omega t}, & x \geq 0. \end{cases} \tag{4.30}$$

Let us now evaluate, as we did in the previous case, the reflected and the transmitted coefficients. First we should note that the transmitted coefficient, which corresponds to the transmitted wave function $\psi_t(x) = Ce^{-k_2'x}$, is zero since $\psi_t(x)$ is a *purely real* function ($\psi_t^*(x) = \psi_t(x)$) and therefore

$$J_{transmitted} = \frac{\hbar}{2im}\left(\psi_t(x)\frac{d\psi_t(x)}{dx} - \psi_t(x)\frac{d\psi_t(x)}{dx}\right) = 0. \tag{4.31}$$

Hence, the reflected coefficient R must be equal to 1. We can obtain this result by applying the continuity conditions at $x = 0$ for (4.17) and (4.29):

$$B = \frac{k_1 - ik'_2}{k_1 + ik'_2} A, \qquad C = \frac{2k_1}{k_1 + ik'_2} A. \tag{4.32}$$

Thus, the reflected coefficient is given by

$$R = \frac{|B|^2}{|A|^2} = \frac{k_1^2 + k'^2_2}{k_1^2 + k'^2_2} = 1. \tag{4.33}$$

We therefore have total reflection, as in the classical case.

There is, however, a difference with the classical case: while none of the particles can be found classically in the region $x > 0$, quantum mechanically there is a *nonzero probability* that the wave function penetrates this *classically forbidden* region. To see this, note that the relative probability density

$$P(x) = |\psi_t(x)|^2 = |C|^2 e^{-2k'_2 x} = \frac{4k_1^2 |A|^2}{k_1^2 + k'^2_2} e^{-2k'_2 x} \tag{4.34}$$

is appreciable near $x = 0$ and falls exponentially to small values as x becomes large; the behavior of the probability density is shown in Figure 4.2.

4.5 The Potential Barrier and Well

Consider a beam of particles of mass m that are sent from the left on a potential barrier

$$V(x) = \begin{cases} 0, & x \leq 0, \\ V_0, & 0 < x < a, \\ 0, & x \geq a. \end{cases} \tag{4.35}$$

This potential, which is repulsive, supports no bound states (Figure 4.3). We are dealing here, as in the case of the potential step, with a one-dimensional *scattering* problem.

Again, let us consider the following two cases which correspond to the particle energies being respectively larger and smaller than the potential barrier.

4.5.1 The Case $E > V_0$

Classically, the particles that approach the barrier from the left at constant momentum, $p_1 = \sqrt{2mE}$, as they enter the region $0 < x < a$ will slow down to a momentum $p_2 = \sqrt{2m(E - V_0)}$. They will maintain the momentum p_2 until they reach the point $x = a$. Then, as soon as they pass beyond the point $x = a$, they will accelerate to a momentum $p_3 = \sqrt{2mE}$ and maintain this value in the entire region $x \geq a$. Since the particles have enough energy to cross the barrier, none of the particles will be reflected back; all the particles will emerge on the right side of $x = a$: *total transmission*.

It is easy to infer the quantum mechanical study from the treatment of the potential step presented in the previous section. We need only to mention that the wave function will display

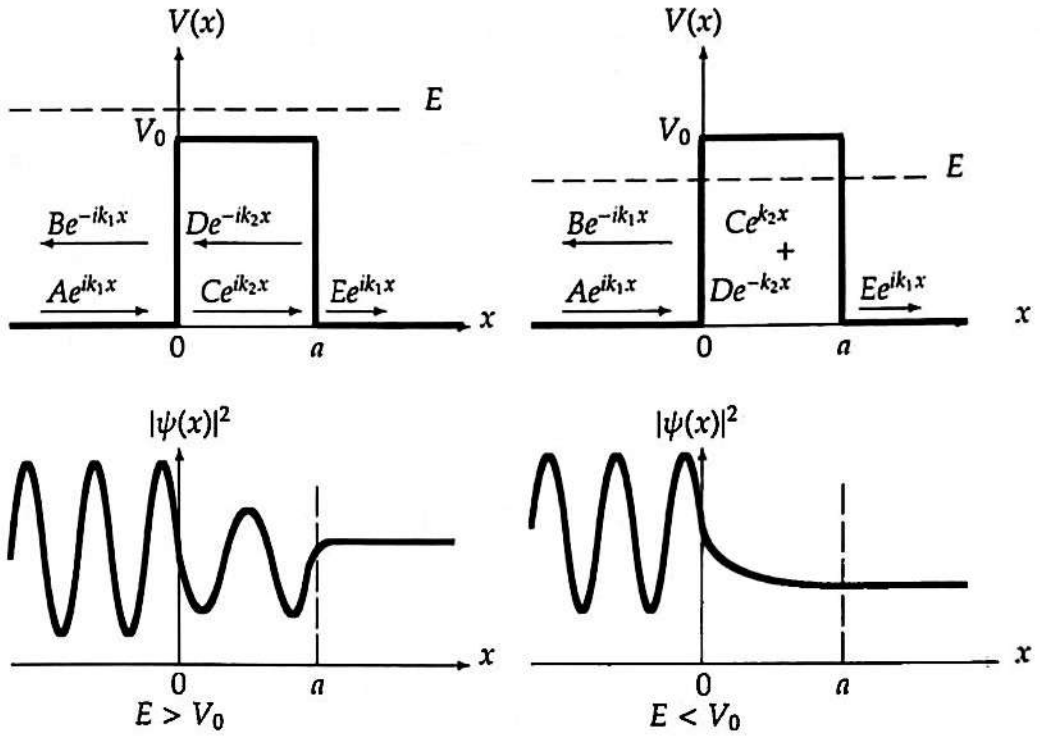


Figure 4.3: Potential barrier and propagation directions of the incident, reflected, and transmitted waves, plus their probability densities $|\psi(x)|^2$ when $E > V_0$ and $E < V_0$.

an oscillatory pattern in all three regions; its amplitude reduces every time the particle enters a new region (see Figure 4.3):

$$\psi(x) = \begin{cases} \psi_1(x) = Ae^{ik_1x} + Be^{-ik_1x}, & x \le 0, \\ \psi_2(x) = Ce^{ik_2x} + De^{-ik_2x}, & 0 < x < a, \\ \psi_3(x) = Ee^{ik_1x}, & x \ge a, \end{cases} \tag{4.36}$$

where $k_1 = \sqrt{2mE/\hbar^2}$ and $k_2 = \sqrt{2m(E - V_0)/\hbar^2}$. The constants B, C, D, and E can be obtained in terms of A from the boundary conditions: $\psi(x)$ and $d\psi/dx$ must be continuous at $x = 0$ and $x = a$, respectively:

$$\psi_1(0) = \psi_2(0), \qquad \frac{d\psi_1(0)}{dx} = \frac{d\psi_2(0)}{dx}, \tag{4.37}$$

$$\psi_2(a) = \psi_3(a), \qquad \frac{d\psi_2(a)}{dx} = \frac{d\psi_3(a)}{dx}. \tag{4.38}$$

These equations yield, respectively

$$A + B = C + D, \tag{4.39}$$

$$ik_1(A - B) = ik_2(C - D) \quad \Longrightarrow \quad A - B = \frac{k_2}{k_1}(C - D), \tag{4.40}$$

$$Ce^{ik_2a} + De^{-ik_2a} = Ee^{ik_1a} \tag{4.41}$$

$$ik_2\left(Ce^{ik_2a} - De^{-ik_2a}\right) = ik_1Ee^{ik_1a} \quad \Longrightarrow \quad Ce^{ik_2a} - De^{-ik_2a} = \frac{k_1}{k_2}Ee^{ik_1a}. \tag{4.42}$$

In what follows, let us find the coefficient E which will be needed in the calculation of the transmission coefficient T. First, adding (4.39) to (4.40), we find

$$2A = C\left(1+\frac{k_2}{k_1}\right)+D\left(1-\frac{k_2}{k_1}\right). \tag{4.43}$$

Similarly, an addition of (4.41) to (4.42) leads to

$$2Ce^{ik_2a} = E\left(1+\frac{k_1}{k_2}\right)e^{ik_1a} \quad \Longrightarrow \quad C = \frac{E}{2}\left(1+\frac{k_1}{k_2}\right)e^{i(k_1-k_2)a}. \tag{4.44}$$

Also, a subtraction of (4.42) from (4.41) yields

$$2De^{-ik_2a} = E\left(1-\frac{k_1}{k_2}\right)e^{ik_1a} \quad \Longrightarrow \quad D = \frac{E}{2}\left(1-\frac{k_1}{k_2}\right)e^{i(k_1+k_2)a}. \tag{4.45}$$

Now, inserting (4.44) and (4.45) into (4.43), we obtain

$$\begin{aligned}
2A &= \frac{E}{2}\left[\left(1+\frac{k_2}{k_1}\right)\left(1+\frac{k_1}{k_2}\right)e^{-ik_2a}+\left(1-\frac{k_2}{k_1}\right)\left(1-\frac{k_1}{k_2}\right)e^{ik_2a}\right]e^{ik_1a} \\
&= \frac{E}{2k_1k_2}\left[(k_1+k_2)^2e^{-ik_2a}-(k_1-k_2)^2e^{ik_2a}\right]e^{ik_1a} \\
&= \frac{E}{2k_1k_2}\left\{\left[(k_1+k_2)^2-(k_1-k_2)^2\right]\cos(k_2a)-i\left[(k_1+k_2)^2+(k_1-k_2)^2\right]\sin(k_2a)\right\}e^{ik_1a} \\
&= \frac{E}{2k_1k_2}\left[4k_1k_2\cos(k_2a)-2i\left(k_1^2+k_2^2\right)\sin(k_2a)\right]e^{ik_1a}.
\end{aligned} \tag{4.46}$$

Thus, the coefficient E is given by

$$E = 4k_1k_2Ae^{-ik_1a}\left[4k_1k_2\cos(k_2a)-2i\left(k_1^2+k_2^2\right)\sin(k_2a)\right]^{-1}, \tag{4.47}$$

which leads at once to the transmission coefficient:

$$\begin{aligned}
T &= \frac{k_1|E|^2}{k_1|A|^2} = \frac{|E|^2}{|A|^2} = \left|\frac{4k_1k_2e^{-ik_1a}}{4k_1k_2\cos(k_2a)-2i\left(k_1^2+k_2^2\right)\sin(k_2a)}\right|^2 \\
&= \frac{(4k_1k_2)^2}{(4k_1k_2)^2\cos^2(k_2a)+4\left(k_1^2+k_2^2\right)^2\sin^2(k_2a)} \\
&= \frac{(4k_1k_2)^2}{(4k_1k_2)^2+\left[-(4k_1k_2)^2+4\left(k_1^2+k_2^2\right)^2\right]\sin^2(k_2a)};
\end{aligned} \tag{4.48}$$

we can reduce this expression to

$$T = \left[1+\frac{1}{4}\left(\frac{k_1^2-k_2^2}{k_1k_2}\right)^2\sin^2(k_2a)\right]^{-1}, \tag{4.49}$$

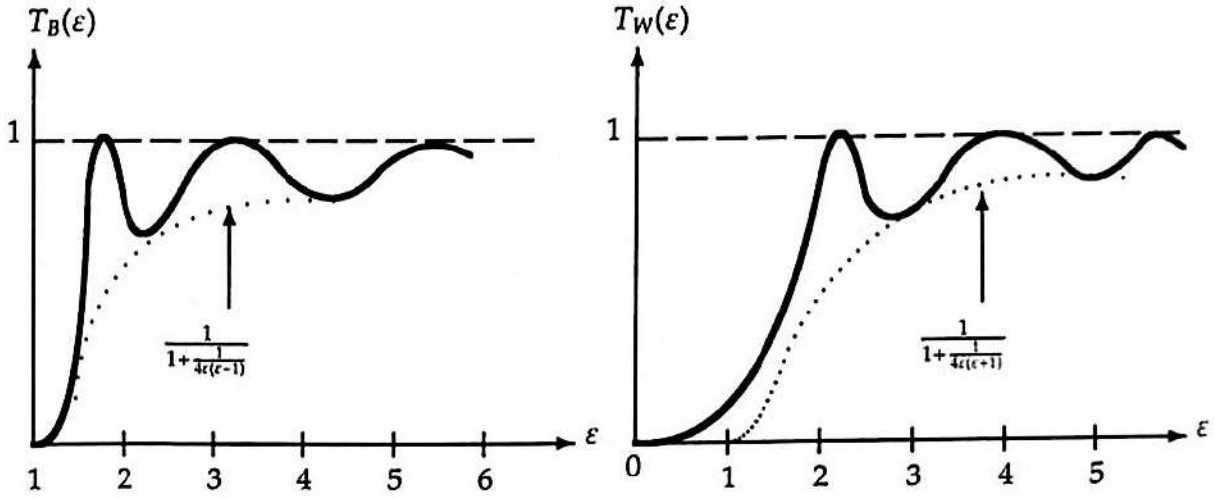


Figure 4.4: Transmission coefficients for a potential barrier, $T_B(\varepsilon) = \frac{4\varepsilon(\varepsilon-1)}{4\varepsilon(\varepsilon-1)+\sin^2(\lambda\sqrt{\varepsilon-1})}$, and for a potential well, $T_W(\varepsilon) = \frac{4\varepsilon(\varepsilon+1)}{4\varepsilon(\varepsilon+1)+\sin^2(\lambda\sqrt{\varepsilon+1})}$.

or to

$$T = \left[1 + \frac{V_0^2}{4E(E-V_0)}\sin^2\left(a\sqrt{2mV_0/\hbar^2}\sqrt{E/V_0-1}\right)\right]^{-1}, \tag{4.50}$$

because

$$\left(\frac{k_1^2-k_2^2}{k_1k_2}\right)^2 = \frac{V_0^2}{E(E-V_0)}. \tag{4.51}$$

Using the notation $\lambda = a\sqrt{2mV_0/\hbar^2}$ and $\varepsilon = E/V_0$, we can rewrite T as

$$T = \left[1 + \frac{1}{4\varepsilon(\varepsilon-1)}\sin^2(\lambda\sqrt{\varepsilon-1})\right]^{-1}. \tag{4.52}$$

Following the same method above that led to the transmission coefficient, we can show that the boundary conditions (4.39)–(4.42) yield the following expression for the reflection coefficient

$$R = \frac{\sin^2(\lambda\sqrt{\varepsilon-1})}{4\varepsilon(\varepsilon-1)+\sin^2(\lambda\sqrt{\varepsilon-1})} = \left[1 + \frac{4\varepsilon(\varepsilon-1)}{\sin^2(\lambda\sqrt{\varepsilon-1})}\right]^{-1}. \tag{4.53}$$

By adding (4.52) and (4.53), we can verify that the reflected and transmission coefficients add up to unity:

$$R+T = \left[1 + \frac{4\varepsilon(\varepsilon-1)}{\sin^2(\lambda\sqrt{\varepsilon-1})}\right]^{-1} + \left[1 + \frac{1}{4\varepsilon(\varepsilon-1)}\sin^2(\lambda\sqrt{\varepsilon-1})\right]^{-1} = 1 \tag{4.54}$$

Special cases

- If $E \gg V_0$, and hence $\varepsilon \gg 1$, the transmission coefficient T becomes asymptotically equal to unity, $T \simeq 1$, and $R \simeq 0$. So, at very high energies and weak potential barrier, the particles would not feel the effect of the barrier; we have total transmission.

- We also have total transmission when $\sin(\lambda\sqrt{\varepsilon-1}) = 0$ or when $\lambda\sqrt{\varepsilon-1} = n\pi$. As shown in Figure 4.4, when $\varepsilon_n = E_n/V_0 = n^2\pi^2\hbar^2/(2ma^2V_0) + 1$ or when the incident energy of the particle is $E_n = V_0 + n^2\pi^2\hbar^2/(2ma^2)$ with $n = 1, 2, 3, \ldots$, we have total transmission, $T(\varepsilon_n) = 1$. The maxima of the transmission coefficient coincide with the energy eigenvalues of the infinite square well potential; these are known as resonances. This resonance phenomenon, which does not occur in classical physics, results from a constructive interference between the incident and the reflected waves. This phenomenon is observed experimentally in a number of cases such as when scattering low-energy ($E \sim 0.1$ eV) electrons off noble atoms (known as the *Ramsauer–Townsend effect*, a consequence of symmetry of noble atoms) and neutrons off nuclei.

- In the limit $\varepsilon \to 1$ we have $\sin(\lambda\sqrt{\varepsilon-1}) \sim \lambda\sqrt{\varepsilon-1}$, hence (4.52) and (4.53) become

$$T = \left(1 + \frac{ma^2V_0}{2\hbar^2}\right)^{-1}, \qquad R = \left(1 + \frac{2\hbar^2}{ma^2V_0}\right)^{-1}. \tag{4.55}$$

The potential well ($V_0 < 0$)
The transmission coefficient (4.52) was derived for the case where $V_0 > 0$, i.e., for a *barrier potential*. Following the same procedure that led to (4.52), we can show that the transmission coefficient for a finite *potential well*, $V_0 < 0$, is given by

$$T_W = \left[1 + \frac{1}{4\varepsilon(\varepsilon+1)}\sin^2(\lambda\sqrt{\varepsilon+1})\right]^{-1}, \tag{4.56}$$

where $\varepsilon = E/|V_0|$ and $\lambda = a\sqrt{2m|V_0|/\hbar^2}$. Notice that there is total transmission whenever $\sin(\lambda\sqrt{\varepsilon+1}) = 0$ or $\lambda\sqrt{\varepsilon+1} = n\pi$. As shown in Figure 4.4, the total transmission, $T_W(\varepsilon_n) = 1$, occurs whenever $\varepsilon_n = E_n/|V_0| = n^2\pi^2\hbar^2/(2ma^2V_0) - 1$ or whenever the incident energy of the particle is $E_n = n^2\pi^2\hbar^2/(2ma^2) - |V_0|$ with $n = 1, 2, 3, \ldots$. We will study in more detail the *symmetric* potential well in Section 4.7.

4.5.2 The Case $E < V_0$: Tunneling

Classically, we would expect total reflection: every particle that arrives at the barrier ($x = 0$) will be reflected back; no particle can penetrate the barrier, where it would have a negative kinetic energy.

We are now going to show that the quantum mechanical predictions differ sharply from their classical counterparts, for the wave function is not zero beyond the barrier. The solutions of the Schrödinger equation in the three regions yield expressions that are similar to (4.36) except that $\psi_2(x) = Ce^{ik_2x} + De^{-ik_2x}$ should be replaced with $\psi_2(x) = Ce^{k_2x} + De^{-k_2x}$:

$$\psi(x) = \begin{cases} \psi_1(x) = Ae^{ik_1x} + Be^{-ik_1x}, & x \le 0, \\ \psi_2(x) = Ce^{k_2x} + De^{-k_2x}, & 0 < x < a, \\ \psi_3(x) = Ee^{ik_1x}, & x \ge a, \end{cases} \tag{4.57}$$

where $k_1^2 = 2mE/\hbar^2$ and $k_2^2 = 2m(V_0 - E)/\hbar^2$. The behavior of the probability density corresponding to this wave function is expected, as displayed in Figure 4.3, to be oscillatory in the regions $x \le 0$ and $x \ge a$, and exponentially decaying for $0 < x < a$.

To find the reflection and transmission coefficients,

$$R = \frac{|B|^2}{|A|^2}, \qquad T = \frac{|E|^2}{|A|^2}, \tag{4.58}$$

we need only to calculate B and E in terms of A. The continuity conditions of the wave function and its derivative at $x = 0$ and $x = a$ yield

$$A + B = C + D, \tag{4.59}$$

$$ik_1(A - B) = k_2(C - D), \tag{4.60}$$

$$Ce^{k_2 a} + De^{-k_2 a} = Ee^{ik_1 a}, \tag{4.61}$$

$$k_2\left(Ce^{k_2 a} - De^{-k_2 a}\right) = ik_1 Ee^{ik_1 a}. \tag{4.62}$$

The last two equations lead to the following expressions for C and D:

$$C = \frac{E}{2}\left(1 + i\frac{k_1}{k_2}\right)e^{(ik_1 - k_2)a}, \qquad D = \frac{E}{2}\left(1 - i\frac{k_1}{k_2}\right)e^{(ik_1 + k_2)a}. \tag{4.63}$$

Inserting these two expressions into the two equations (4.59) and (4.60) and dividing by A, we can show that these two equations reduce, respectively, to

$$1 + \frac{B}{A} = \frac{E}{A}e^{ik_1 a}\left[\cosh(k_2 a) - i\frac{k_1}{k_2}\sinh(k_2 a)\right], \tag{4.64}$$

$$1 - \frac{B}{A} = \frac{E}{A}e^{ik_1 a}\left[\cosh(k_2 a) + i\frac{k_2}{k_1}\sinh(k_2 a)\right]. \tag{4.65}$$

Solving these two equations for B/A and E/A, we obtain

$$\frac{B}{A} = -i\frac{k_1^2 + k_2^2}{k_1 k_2}\sinh(k_2 a)\left[2\cosh(k_2 a) + i\frac{k_2^2 - k_1^2}{k_1 k_2}\sinh(k_2 a)\right]^{-1}, \tag{4.66}$$

$$\frac{E}{A} = 2e^{-ik_1 a}\left[2\cosh(k_2 a) + i\frac{k_2^2 - k_1^2}{k_1 k_2}\sinh(k_2 a)\right]^{-1}. \tag{4.67}$$

Thus, the coefficients R and T become

$$R = \frac{|B|^2}{|A|^2} = \left(\frac{k_1^2 + k_2^2}{k_1 k_2}\right)^2 \sinh^2(k_2 a)\left[4\cosh^2(k_2 a) + \left(\frac{k_2^2 - k_1^2}{k_1 k_2}\right)^2 \sinh^2(k_2 a)\right]^{-1}, \tag{4.68}$$

$$T = \frac{|E|^2}{|A|^2} = 4\left[4\cosh^2(k_2 a) + \left(\frac{k_2^2 - k_1^2}{k_1 k_2}\right)^2 \sinh^2(k_2 a)\right]^{-1}. \tag{4.69}$$

Since $\cosh^2(k_2 a) = 1 + \sinh^2(k_2 a)$, we can reduce (4.68) to

$$R = \frac{1}{4}T\left(\frac{k_1^2 + k_2^2}{k_1 k_2}\right)^2 \sinh^2(k_2 a) = \left[1 + 4\left(\frac{k_1 k_2}{k_1^2 + k_2^2}\right)^2 \frac{1}{\sinh^2(k_2 a)}\right]^{-1}, \tag{4.70}$$

and (4.69) to

$$T = \left[1 + \frac{1}{4}\left(\frac{k_1^2 + k_2^2}{k_1 k_2}\right)^2 \sinh^2(k_2 a)\right]^{-1}. \tag{4.71}$$

Note that T is *finite*. This means that the probability for the transmission of the particles into the region $x \geq a$ is *not zero* (in classical physics, however, the particle can in no way make it into the $x \geq 0$ region). This is a purely quantum mechanical effect which is due to the *wave aspect* of microscopic objects; it is known as the *tunneling effect*: *quantum mechanical objects can tunnel through classically impenetrable barriers*. This *barrier penetration* effect has important applications in various branches of modern physics ranging from particle and nuclear physics to semiconductor devices. For instance, radioactive decays and charge transport in electronic devices are typical examples of the tunneling effect.

Now since

$$\left(\frac{k_1^2 + k_2^2}{k_1 k_2}\right)^2 = \left(\frac{V_0}{\sqrt{E(V_0 - E)}}\right)^2 = \frac{V_0^2}{E(V_0 - E)}, \tag{4.72}$$

we can rewrite (4.70) and (4.71) as follows:

$$R = \frac{1}{4}\frac{V_0^2 T}{E(V_0 - E)} \sinh^2\left(\frac{a}{\hbar}\sqrt{2m(V_0 - E)}\right), \tag{4.73}$$

$$T = \left[1 + \frac{1}{4}\frac{V_0^2}{E(V_0 - E)} \sinh^2\left(\frac{a}{\hbar}\sqrt{2m(V_0 - E)}\right)\right]^{-1}, \tag{4.74}$$

or

$$R = \frac{T}{4\varepsilon(1 - \varepsilon)} \sinh^2\left(\lambda\sqrt{1 - \varepsilon}\right), \tag{4.75}$$

$$T = \left[1 + \frac{1}{4\varepsilon(1 - \varepsilon)} \sinh^2\left(\lambda\sqrt{1 - \varepsilon}\right)\right]^{-1}, \tag{4.76}$$

where $\lambda = a\sqrt{2mV_0/\hbar^2}$ and $\varepsilon = E/V_0$.

Special cases

- If $E \ll V_0$, hence $\varepsilon \ll 1$ or $\lambda\sqrt{1 - \varepsilon} \gg 1$, we may approximate $\sinh\left(\lambda\sqrt{1 - \varepsilon}\right) \simeq \frac{1}{2}\exp\left(\lambda\sqrt{1 - \varepsilon}\right)$. We can thus show that the transmission coefficient (4.76) becomes asymptotically equal to

$$\begin{aligned} T &\simeq \left\{\frac{1}{4\varepsilon(1 - \varepsilon)}\left[\frac{1}{2}e^{\lambda\sqrt{1-\varepsilon}}\right]^2\right\}^{-1} = 16\varepsilon(1 - \varepsilon)e^{-2\lambda\sqrt{1-\varepsilon}} \\ &= \frac{16E}{V_0}\left(1 - \frac{E}{V_0}\right)e^{-(2a/\hbar)\sqrt{2m(V_0-E)}}. \end{aligned} \tag{4.77}$$

This shows that the transmission coefficient is not zero, as it would be classically, but has a finite value. So, quantum mechanically, there is a finite tunneling beyond the barrier, $x > a$.

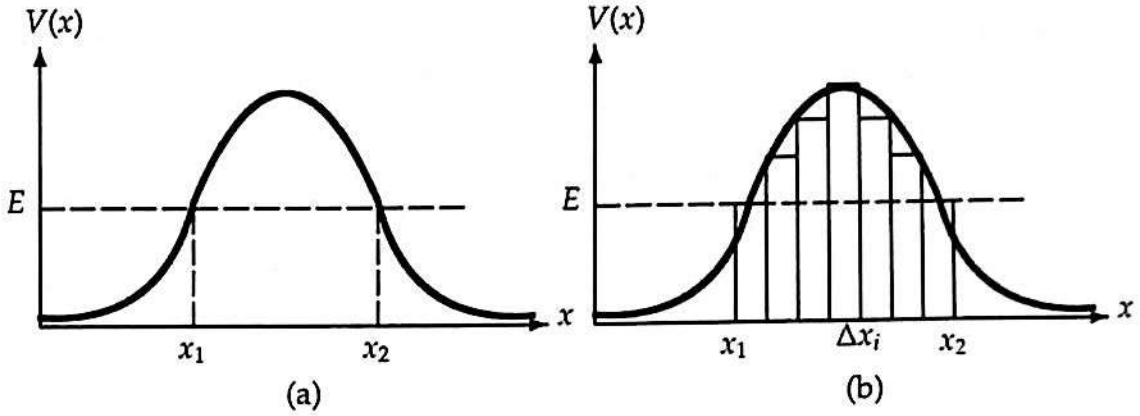


Figure 4.5: (a) Tunneling though a potential barrier. (b) Approximation of a smoothly varying potential $V(x)$ by square barriers.

- When $E \simeq V_0$, hence $\varepsilon \simeq 1$, we can verify that (4.75) and (4.76) lead to the relations (4.55).
- Taking the classical limit $\hbar \to 0$, the coefficients (4.75) and (4.76) reduce to the classical result: $R \to 1$ and $T \to 0$.

4.5.3 The Tunneling Effect

In general, the tunneling effect consists of the propagation of a particle through a region where the particle's energy is smaller than the potential energy $E < V(x)$. Classically this region, defined by $x_1 < x < x_2$ (Figure 4.5a), is forbidden to the particle where its kinetic energy would be negative; the points $x = x_1$ and $x = x_2$ are known as the *classical turning points*. Quantum mechanically, however, since particles display wave features, the quantum waves can tunnel through the barrier.

As shown in the square barrier example, the particle has a finite probability of tunneling through the barrier. In this case we managed to find an analytical expression (4.76) for the tunneling probability only because we dealt with a simple square potential. Analytic expressions cannot be obtained for potentials with arbitrary spatial dependence. In such cases one needs approximations. The Wentzel–Kramers–Brillouin (WKB) method (Chapter 9) provides one of the most useful approximation methods. We will show that the transmission coefficient for a barrier potential $V(x)$ is given by

$$T \sim \exp\left\{-\frac{2}{\hbar}\int_{x_1}^{x_2} dx\,\sqrt{2m\,[V(x)-E]}\right\}. \tag{4.78}$$

We can obtain this relation by means of a crude approximation. For this, we need simply to take the classically forbidden region $x_1 < x < x_2$ (Figure 4.5b) and divide it into a series of small intervals Δx_i. If Δx_i is small enough, we may approximate the potential $V(x_i)$ at each point x_i by a square potential barrier. Thus, we can use (4.77) to calculate the transmission probability corresponding to $V(x_i)$:

$$T_i \sim \exp\left[-\frac{2\Delta x_i}{\hbar}\sqrt{2m(V(x_i)-E)}\right]. \tag{4.79}$$

The transmission probability for the general potential of Figure 4.5, where we divided the region $x_1 < x < x_2$ into a very large number of small intervals Δx_i, is given by

$$\begin{aligned} T &\sim \lim_{N\to\infty} \prod_{i=1}^{N} \exp\left[-\frac{2\Delta x_i}{\hbar}\sqrt{2m(V(x_i)-E)}\right] \\ &= \exp\left[-\frac{2}{\hbar}\lim_{\Delta x_i\to 0}\sum_i \Delta x_i \sqrt{2m(V(x_i)-E)}\right] \\ &\longrightarrow \exp\left[-\frac{2}{\hbar}\int_{x_1}^{x_2} dx\,\sqrt{2m\,[V(x)-E]}\right]. \end{aligned} \tag{4.80}$$

The approximation leading to this relation is valid, as will be shown in Chapter 9, only if the potential $V(x)$ is a smooth, slowly varying function of x.

4.6 The Infinite Square Well Potential

4.6.1 The Asymmetric Square Well

Consider a particle of mass m confined to move inside an infinitely deep asymmetric potential well

$$V(x) = \begin{cases} +\infty, & x < 0, \\ 0, & 0 \le x \le a, \\ +\infty, & x > a. \end{cases} \tag{4.81}$$

Classically, the particle remains confined inside the well, moving at constant momentum $p = \pm\sqrt{2mE}$ back and forth as a result of repeated reflections from the walls of the well.

Quantum mechanically, we expect this particle to have only bound state solutions and a discrete nondegenerate energy spectrum. Since $V(x)$ is infinite outside the region $0 \le x \le a$, the wave function of the particle must be zero outside the boundary. Hence we can look for solutions only inside the well

$$\frac{d^2\psi(x)}{dx^2} + k^2\psi(x) = 0, \tag{4.82}$$

with $k^2 = 2mE/\hbar^2$; the solutions are

$$\psi(x) = A'e^{ikx} + B'e^{-ikx} \quad\Longrightarrow\quad \psi(x) = A\sin(kx) + B\cos(kx). \tag{4.83}$$

The wave function vanishes at the walls, $\psi(0) = \psi(a) = 0$: the condition $\psi(0) = 0$ gives $B = 0$, while $\psi(a) = A\sin(ka) = 0$ gives

$$k_n a = n\pi \qquad (n = 1, 2, 3, \cdots). \tag{4.84}$$

This condition determines the energy

$$\boxed{E_n = \frac{\hbar^2}{2m}k_n^2 = \frac{\hbar^2\pi^2}{2ma^2}n^2 = n^2E_1, \qquad (n = 1, 2, 3, \cdots).} \tag{4.85}$$

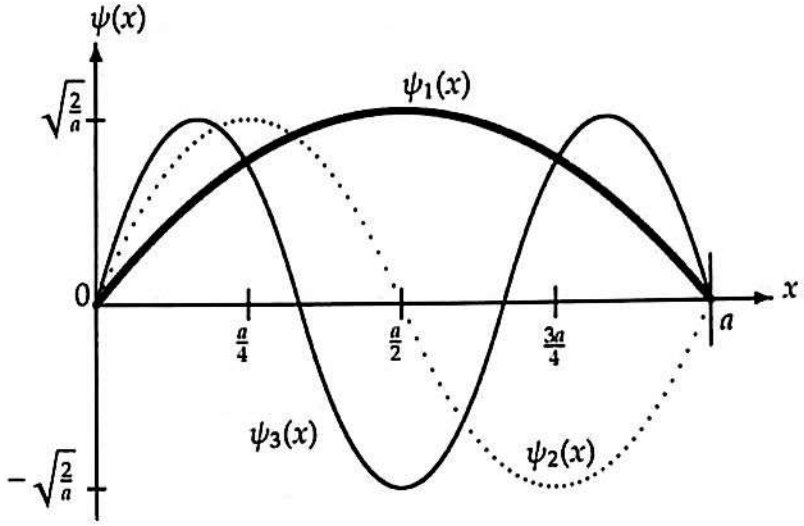


Figure 4.6: Three lowest states of an infinite potential well, $\psi_n(x) = \sqrt{2/a}\sin(n\pi x/a)$; the states $\psi_{2n+1}(x)$ and $\psi_{2n}(x)$ are even and odd, respectively, with respect to $x = a/2$.

The energy is *quantized*; only certain values are permitted. This is expected since the states of a particle which is confined to a limited region of space are *bound states* and the energy spectrum is *discrete*. This is in sharp contrast to classical physics where the energy of the particle, given by $E = p^2/(2m)$, takes any value; the classical energy evolves *continuously*.

As it can be inferred from (4.85), we should note that the energy between adjacent levels is not constant:

$$E_{n+1} - E_n = (2n+1)E_1, \tag{4.86}$$

which leads to

$$\frac{E_{n+1} - E_n}{E_n} = \frac{(n+1)^2 - n^2}{n^2} = \frac{2n+1}{n^2}. \tag{4.87}$$

In the classical limit $n \to \infty$,

$$\lim_{n\to\infty} \frac{E_{n+1} - E_n}{E_n} = \lim_{n\to\infty} \frac{2n+1}{n^2} = 0, \tag{4.88}$$

the levels become so close together as to be practically indistinguishable.

Since $B = 0$ and $k_n = n\pi/a$, (4.83) yields $\psi_n(x) = A\sin(n\pi x/a)$. We can choose the constant A so that $\psi_n(x)$ is normalized:

$$1 = \int_0^a |\psi_n(x)|^2 dx = |A|^2 \int_0^a \sin^2\left(\frac{n\pi}{a}x\right) dx \quad \Longrightarrow \quad A = \sqrt{\frac{2}{a}}, \tag{4.89}$$

hence

$$\boxed{\psi_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right) \qquad (n = 1,2,3,\cdots).} \tag{4.90}$$

The first few functions are plotted in Figure 4.6.

The solution of the time-independent Schrödinger equation has thus given us the energy (4.85) and the wave function (4.90). There is then an infinite sequence of discrete energy levels

corresponding to the positive integer values of the *quantum number* n. It is clear that $n = 0$ yields an uninteresting result: $\psi_0(x) = 0$ and $E_0 = 0$; later, we will examine in more detail the physical implications of this. So, the lowest energy, or *ground state* energy, corresponds to $n = 1$; it is $E_1 = \hbar^2\pi^2/(2ma^2)$. As will be explained later, this is called the *zero-point energy*, for there exists no state with zero energy. The states corresponding to $n = 2, 3, 4, \ldots$ are called *excited states*; their energies are given by $E_n = n^2 E_1$. As mentioned in Theorem 4.2, each function $\psi_n(x)$ has $(n-1)$ nodes. Figure 4.6 shows that the functions $\psi_{2n+1}(x)$ are even and the functions $\psi_{2n}(x)$ are odd with respect to the center of the well; we will study this in Section 4.6.2 when we consider the symmetric potential well. Note that none of the energy levels is degenerate (there is only one eigenfunction for each energy level) and that the wave functions corresponding to different energy levels are orthogonal:

$$\int_0^a \psi_m^*(x)\psi_n(x)\,dx = \delta_{mn}. \tag{4.91}$$

Since we are dealing with stationary states and since $E_n = n^2 E_1$, the most general solutions of the time-dependent Schrödinger equation are given by

$$\Psi(x,t) = \sum_{n=1}^{\infty} \psi_n(x) e^{-iE_n t/\hbar} = \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{a}\right) e^{-in^2 E_1 t/\hbar}. \tag{4.92}$$

Zero-point energy

Let us examine why there is no state with zero energy for a square well potential. If the particle has zero energy, it will be at rest inside the well, and this violates Heisenberg's uncertainty principle. By *localizing* or confining the particle to a limited region in space, it will acquire a *finite* momentum leading to a minimum kinetic energy. That is, the localization of the particle's motion to $0 \le x \le a$ implies a position uncertainty of order $\Delta x \sim a$ which, according to the uncertainty principle, leads to a minimum momentum uncertainty $\Delta p \sim \hbar/a$ and this in turn leads to a minimum kinetic energy of order $\hbar^2/(2ma^2)$. This is in qualitative agreement with the exact value $E_1 = \pi^2\hbar^2/(2ma^2)$. In fact, as will be shown in (4.229), an accurate evaluation of Δp_1 leads to a zero-point energy which is equal to E_1.

Note that, as the momentum uncertainty is inversely proportional to the width of the well, $\Delta p \sim \hbar/a$, if the width decreases (i.e., the particle's position is confined further and further), the uncertainty on $\hat{P}$ will increase. This makes the particle move faster and faster, so the zero-point energy will also increase. Conversely, if the width of the well increases, the zero-point energy decreases, but it will never vanish.

The zero-point energy therefore reflects the necessity of a *minimum motion* of a particle due to localization. The zero-point energy occurs in all bound state potentials. In the case of binding potentials, the lowest energy state has an energy which is higher than the minimum of the potential energy. This is in sharp contrast to classical mechanics, where the lowest possible energy is equal to the minimum value of the potential energy, with zero kinetic energy. In quantum mechanics, however, the lowest state does not minimize the potential alone, but applies to the sum of the kinetic and potential energies, and this leads to a finite ground state or zero-point energy. This concept has far-reaching physical consequences in the realm of the microscopic world. For instance, without the zero-point motion, atoms would not be stable, for the electrons would fall into the nuclei. Also, it is the zero-point energy which prevents helium from freezing at very low temperatures.

The following example shows that the zero-point energy is also present in macroscopic systems, but it is infinitesimally small. In the case of microscopic systems, however, it has a nonnegligible size.

Example 4.1 (Zero-point energy)
To illustrate the idea that the zero-point energy gets larger by going from macroscopic to microscopic systems, calculate the zero-point energy for a particle in an infinite potential well for the following three cases:

(a) a 100 g ball confined on a 5 m long line,
(b) an oxygen atom confined to a 2×10^{-10} m lattice, and
(c) an electron confined to a 10^{-10} m atom.

Solution

(a) The zero-point energy of a 100 g ball that is confined to a 5 m long line is

$$E = \frac{\hbar^2\pi^2}{2ma^2} \simeq \frac{10 \times 10^{-68}\,\mathrm{J}}{2 \times 0.1 \times 25} \simeq 2 \times 10^{-68}\,\mathrm{J} = 1.25 \times 10^{-49}\,\mathrm{eV}. \tag{4.93}$$

This energy is too small to be detected, much less measured, by any known experimental technique.

(b) For the zero-point energy of an oxygen atom confined to a 2×10^{-10} m lattice, since the oxygen atom has 16 nucleons, its mass is of the order of $m \simeq 16 \times 1.6 \times 10^{-27}$ kg $\simeq 26 \times 10^{-27}$ kg, so we have

$$E = \frac{10^{-67}\,\mathrm{J}}{2 \times 26 \times 10^{-27} \times 4 \times 10^{-20}} \simeq 0.5 \times 10^{-22}\,\mathrm{J} \simeq 3 \times 10^{-4}\,\mathrm{eV}. \tag{4.94}$$

(c) The zero-point energy of an electron ($m \sim 10^{-30}$ kg) that is confined to an atom ($a \sim 1 \times 10^{-10}$ m) is

$$E = \frac{10^{-67}\,\mathrm{J}}{2 \times 10^{-30} \times 10^{-20}} \simeq 5 \times 10^{-18}\,\mathrm{J} \simeq 30\,\mathrm{eV}. \tag{4.95}$$

This energy is important at the atomic scale, for the binding energy of a hydrogen electron is about 14 eV. So the zero-point energy is negligible for macroscopic objects, but important for microscopic systems.

4.6.2 The Symmetric Potential Well

What happens if the potential (4.81) is translated to the left by a distance of $a/2$ to become symmetric?

$$V(x) = \begin{cases} +\infty, & x < -a/2, \\ 0, & -a/2 \le x \le a/2, \\ +\infty, & x > a/2. \end{cases} \tag{4.96}$$

First, we would expect the energy spectrum (4.85) to remain unaffected by this translation, since the Hamiltonian is invariant under spatial translations; as it contains only a kinetic part, it commutes with the particle's momentum, $[\hat{H}, \hat{P}] = 0$. The energy spectrum is discrete and nondegenerate.

Second, earlier in this chapter we saw that for symmetric potentials, $V(-x) = V(x)$, the wave function of bound states must be either even or odd. The wave function corresponding to the potential (4.96) can be written as follows:

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left[\frac{n\pi}{a}\left(x + \frac{a}{2}\right)\right] = \begin{cases} \sqrt{\frac{2}{a}} \cos(\frac{n\pi}{a}x) & (n = 1,3,5,7,\cdots), \\ \sqrt{\frac{2}{a}} \sin(\frac{n\pi}{a}x) & (n = 2,4,6,8,\cdots). \end{cases} \tag{4.97}$$

That is, the wave functions corresponding to odd quantum numbers $n = 1,3,5,\ldots$ are symmetric, $\psi(-x) = \psi(x)$, and those corresponding to even numbers $n = 2,4,6,\ldots$ are antisymmetric, $\psi(-x) = -\psi(x)$.

4.7 The Finite Square Well Potential

Consider a particle of mass m moving in the following symmetric potential:

$$V(x) = \begin{cases} V_0, & x < -a/2, \\ 0, & -a/2 \le x \le a/2, \\ V_0, & x > a/2. \end{cases} \tag{4.98}$$

The two physically interesting cases are $E > V_0$ and $E < V_0$ (see Figure 4.7). We expect the solutions to yield a *continuous* doubly-degenerate energy spectrum for $E > V_0$ and a *discrete* nondegenerate spectrum for $0 < E < V_0$.

4.7.1 The Scattering Solutions ($E > V_0$)

Classically, if the particle is initially incident from left with constant momentum $\sqrt{2m(E - V_0)}$, it will speed up to $\sqrt{2mE}$ between $-a/2 \le x \le a/2$ and then slow down to its initial momentum in the region $x > a/2$. All the particles that come from the left will be transmitted, none will be reflected back; therefore $T = 1$ and $R = 0$.

Quantum mechanically, and as we did for the step and barrier potentials, we can verify that we get a *finite* reflection coefficient. The solution is straightforward to obtain; just follow the procedure outlined in the previous two sections. The wave function has an oscillating pattern in all three regions (see Figure 4.7).

4.7.2 The Bound State Solutions ($0 < E < V_0$)

Classically, when $E < V_0$ the particle is completely confined to the region $-a/2 \le x \le a/2$; it will bounce back and forth between $x = -a/2$ and $x = a/2$ with constant momentum $p = \sqrt{2mE}$.

Quantum mechanically, the solutions are particularly interesting for they are expected to yield a *discrete* energy spectrum and wave functions that decay in the two regions $x < -a/2$ and $x > a/2$, but oscillate in $-a/2 \le x \le a/2$. In these three regions, the Schrödinger equation

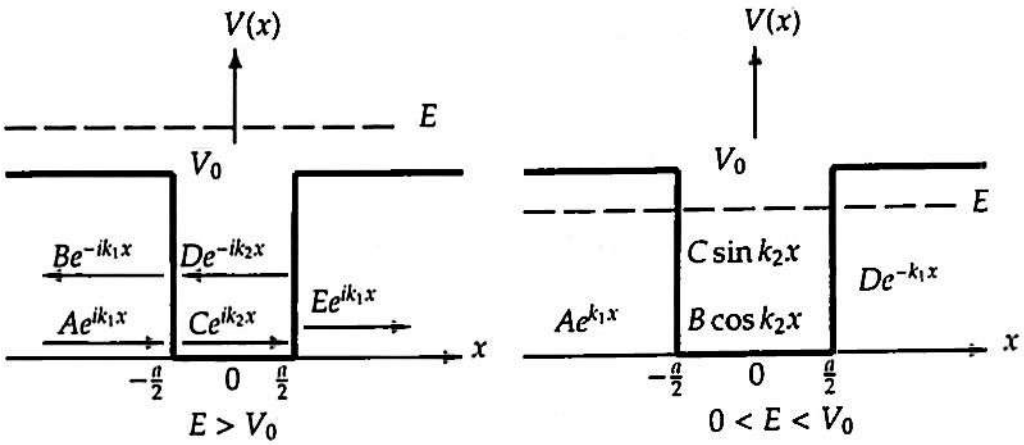


Figure 4.7: Finite square well potential and propagation directions of the incident, reflected and transmitted waves when $E > V_0$ and $0 < E < V_0$.

can be written as

$$\left(\frac{d^2}{dx^2} - k_1^2\right)\psi_1(x) = 0 \qquad \left(x < -\frac{1}{2}a\right), \tag{4.99}$$

$$\left(\frac{d^2}{dx^2} + k_2^2\right)\psi_2(x) = 0 \qquad \left(-\frac{1}{2}a \le x \le \frac{1}{2}a\right), \tag{4.100}$$

$$\left(\frac{d^2}{dx^2} - k_1^2\right)\psi_3(x) = 0 \qquad \left(x > \frac{1}{2}a\right), \tag{4.101}$$

where $k_1^2 = 2m(V_0 - E)/\hbar^2$ and $k_2^2 = 2mE/\hbar^2$. Eliminating the physically unacceptable solutions which grow exponentially for large values of $|x|$, we can write the solution to this Schrödinger equation in the regions $x < -a/2$ and $x > a/2$ as follows:

$$\psi_1(x) = Ae^{k_1 x} \qquad \left(x < -\frac{1}{2}a\right), \tag{4.102}$$

$$\psi_3(x) = De^{-k_1 x} \qquad \left(x > \frac{1}{2}a\right). \tag{4.103}$$

As mentioned in (4.4), since the bound state eigenfunctions of symmetric one-dimensional Hamiltonians are either even or odd under space inversion, the solutions of (4.99) to (4.101) are then either antisymmetric (odd)

$$\psi_a(x) = \begin{cases} Ae^{k_1 x}, & x < -a/2, \\ C\sin(k_2 x), & -a/2 \le x \le a/2, \\ De^{-k_1 x}, & x > a/2, \end{cases} \tag{4.104}$$

or symmetric (even)

$$\psi_s(x) = \begin{cases} Ae^{k_1 x}, & x < -a/2, \\ B\cos(k_2 x), & -a/2 \le x \le a/2, \\ De^{-k_1 x}, & x > a/2. \end{cases} \tag{4.105}$$

To determine the eigenvalues, we need to use the continuity conditions at $x = \pm a/2$. The continuity of the logarithmic derivative, $(1/\psi_a(x))d\psi_a(x)/dx$, of $\psi_a(x)$ at $x = \pm a/2$ yields

$$k_2 \cot\left(\frac{k_2 a}{2}\right) = -k_1. \tag{4.106}$$

Similarly, the continuity of $(1/\psi_s(x))d\psi_s(x)/dx$ at $x = \pm a/2$ gives

$$k_2 \tan\left(\frac{k_2 a}{2}\right) = k_1. \tag{4.107}$$

The transcendental equations (4.106) and (4.107) cannot be solved directly; we can solve them either graphically or numerically. To solve these equations graphically, we need only to rewrite them in the following suggestive forms:

$$-\alpha_n \cot \alpha_n = \sqrt{R^2 - \alpha_n^2} \qquad \text{(for odd states)}, \tag{4.108}$$

$$\alpha_n \tan \alpha_n = \sqrt{R^2 - \alpha_n^2} \qquad \text{(for even states)}, \tag{4.109}$$

where $\alpha_n^2 = (k_2 a/2)^2 = ma^2 E_n/(2\hbar^2)$ and $R^2 = ma^2 V_0/(2\hbar^2)$; these equations are obtained by inserting $k_1 = \sqrt{2m(V_0 - E)/\hbar^2}$ and $k_2 = \sqrt{2mE/\hbar^2}$ into (4.106) and (4.107). The left-hand sides of (4.108) and (4.109) consist of trigonometric functions; the right-hand sides consist of a circle of radius R. The solutions are given by the points where the circle $\sqrt{R^2 - \alpha_n^2}$ intersects the functions $-\alpha_n \cot \alpha_n$ and $\alpha_n \tan \alpha_n$ (Figure 4.8). The solutions form a *discrete* set. As illustrated in Figure 4.8, the intersection of the small circle with the curve $\alpha_n \tan \alpha_n$ yields only one bound state, $n = 0$, whereas the intersection of the larger circle with $\alpha_n \tan \alpha_n$ yields two bound states, $n = 0, 2$, and its intersection with $-\alpha_n \cot \alpha_n$ yields two other bound states, $n = 1, 3$.

The number of solutions depends on the size of R, which in turn depends on the depth V_0 and the width a of the well, since $R = \sqrt{ma^2 V_0/(2\hbar^2)}$. The deeper and broader the well, the larger the value of R, and hence the greater the number of bound states. Note that there is always at least one bound state (i.e., one intersection) no matter how small V_0 is. When

$$0 < R < \frac{\pi}{2} \qquad \text{or} \qquad 0 < V_0 < \left(\frac{\pi}{2}\right)^2 \frac{2\hbar^2}{ma^2}, \tag{4.110}$$

there is only one bound state corresponding to $n = 0$ (see Figure 4.8); this state—the ground state—is even. Then, and when

$$\frac{\pi}{2} < R < \pi \qquad \text{or} \qquad \left(\frac{\pi}{2}\right)^2 \frac{2\hbar^2}{ma^2} < V_0 < \pi^2 \frac{2\hbar^2}{ma^2}, \tag{4.111}$$

there are two bound states: an even state (the ground state) corresponding to $n = 0$ and the first odd state corresponding to $n = 1$. Now, if

$$\pi < R < \frac{3\pi}{2} \qquad \text{or} \qquad \pi^2 \frac{2\hbar^2}{ma^2} < V_0 < \left(\frac{3\pi}{2}\right)^2 \frac{2\hbar^2}{ma^2}, \tag{4.112}$$

there exist three bound states: the ground state (even state), $n = 0$, the first excited state (odd state), corresponding to $n = 1$, and the second excited state (even state), which corresponds

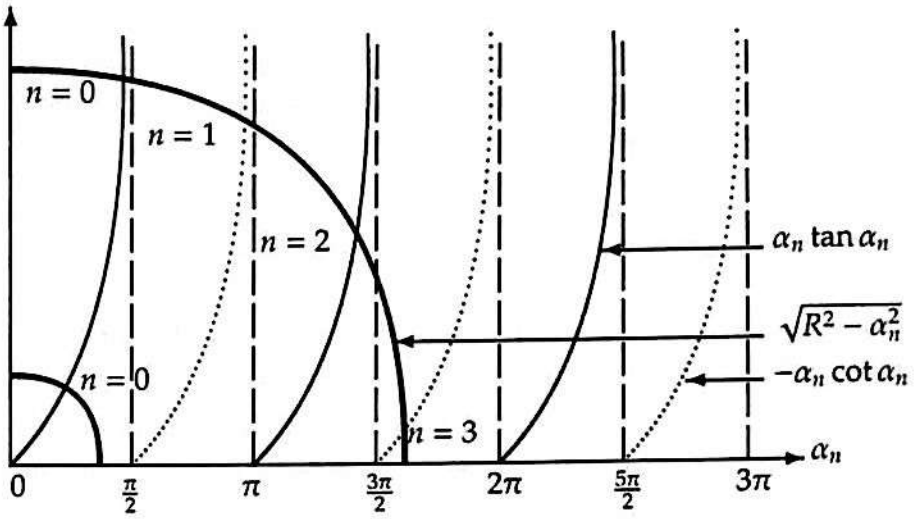


Figure 4.8: Graphical solutions for the finite square well potential: they are given by the intersections of $\sqrt{R^2 - \alpha_n^2}$ with $\alpha_n \tan \alpha_n$ and $-\alpha_n \cot \alpha_n$, where $\alpha_n^2 = ma^2 E_n/(2\hbar^2)$ and $R^2 = ma^2 V_0/(2\hbar^2)$.

to $n = 2$. In general, the well's depth at which n states are allowed is given by

$$R = \frac{n\pi}{2} \qquad \text{or} \qquad V_0 = \left(\frac{\pi}{2}\right)^2 \frac{2\hbar^2}{ma^2} n^2. \tag{4.113}$$

The spectrum, therefore, consists of a set of alternating even and odd states: the lowest state, the ground state, is even, the next state (first excited state) is odd, and so on.

In the limiting case $V_0 \to \infty$, the circle's radius R also becomes infinite, and hence the function $\sqrt{R^2 - \alpha_n^2}$ will cross $-\alpha_n \cot \alpha_n$ and $\alpha_n \tan \alpha_n$ at the asymptotes $\alpha_n = n\pi/2$, because when $V_0 \to \infty$ both $\tan \alpha_n$ and $\cot \alpha_n$ become infinite:

$$\tan \alpha_n \to \infty \implies \alpha_n = \frac{2n+1}{2}\pi \qquad (n = 0, 1, 2, 3, \cdots), \tag{4.114}$$

$$\cot \alpha_n \to \infty \implies \alpha_n = n\pi \qquad (n = 1, 2, 3, \cdots). \tag{4.115}$$

Combining these two cases, we obtain

$$\alpha_n = \frac{n\pi}{2} \qquad (1, 2, 3, \ldots). \tag{4.116}$$

Since $\alpha_n^2 = ma^2 E_n/(2\hbar^2)$ we see that we recover the energy expression for the infinite well:

$$\alpha_n = \frac{n\pi}{2} \longrightarrow E_n = \frac{\pi^2 \hbar^2}{2ma^2} n^2. \tag{4.117}$$

Example 4.2

Find the number of bound states and the corresponding energies for the finite square well potential when: (a) $R = 1$ (i.e., $\sqrt{ma^2 V_0/(2\hbar^2)} = 1$), and (b) $R = 2$.

Solution

(a) From Figure 4.8, when $R = \sqrt{ma^2V_0/(2\hbar^2)} = 1$, there is only one bound state since $\alpha_n \leq R$. This bound state corresponds to $n = 0$. The corresponding energy is given by the intersection of $\alpha_0 \tan\alpha_0$ with $\sqrt{1-\alpha_0^2}$:

$$\alpha_0 \tan\alpha_0 = \sqrt{1-\alpha_0^2} \implies \alpha_0^2(1+\tan^2\alpha_0) = 1 \implies \cos^2\alpha_0 = \alpha_0^2. \tag{4.118}$$

The solution of $\cos^2\alpha_0 = \alpha_0^2$ is given numerically by $\alpha_0 = 0.739\,09$. Thus, the corresponding energy is determined by the relation $\sqrt{ma^2E_0/(2\hbar^2)} = 0.739\,09$, which yields $E_0 \simeq 1.1\hbar^2/(ma^2)$.

(b) When $R = 2$ there are two bound states resulting from the intersections of $\sqrt{4-\alpha_0^2}$ with $\alpha_0 \tan\alpha_0$ and $-\alpha_1 \cot\alpha_1$; they correspond to $n = 0$ and $n = 1$, respectively. The numerical solutions of the corresponding equations

$$\alpha_0 \tan\alpha_0 = \sqrt{4-\alpha_0^2} \implies 4\cos^2\alpha_0 = \alpha_0^2, \tag{4.119}$$

$$-\alpha_1 \cot\alpha_1 = \sqrt{4-\alpha_1^2} \implies 4\sin^2\alpha_1 = \alpha_1^2, \tag{4.120}$$

yield $\alpha_0 \simeq 1.03$ and $\alpha_1 \simeq 1.9$, respectively. The corresponding energies are

$$\alpha_0 = \sqrt{\frac{ma^2E_0}{2\hbar^2}} \simeq 1.03 \implies E_0 \simeq \frac{2.12\hbar^2}{ma^2}, \tag{4.121}$$

$$\alpha_1 = \sqrt{\frac{ma^2E_1}{2\hbar^2}} \simeq 1.9 \implies E_1 \simeq \frac{7.22\hbar^2}{ma^2}. \tag{4.122}$$

4.8 The Harmonic Oscillator

The harmonic oscillator is one of those few problems that are important to all branches of physics. It provides a useful model for a variety of vibrational phenomena that are encountered, for instance, in classical mechanics, electrodynamics, statistical mechanics, solid state, atomic, nuclear, and particle physics. In quantum mechanics, it serves as an invaluable tool to illustrate the basic concepts and the formalism.

The Hamiltonian of a particle of mass m which oscillates with an angular frequency ω under the influence of a one-dimensional harmonic potential is

$$\hat{H} = \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2\hat{X}^2. \tag{4.123}$$

The problem is how to find the energy eigenvalues and eigenstates of this Hamiltonian. This problem can be studied by means of two separate methods. The first method, called the *analytic method*, consists in solving the time-independent Schrödinger equation (TISE) for the Hamiltonian (4.123). The second method, called the *ladder* or *algebraic method*, does not deal

with solving the Schrödinger equation, but deals instead with operator algebra involving operators known as the *creation* and *annihilation* or *ladder* operators; this method is in essence a matrix formulation, because it expresses the various quantities in terms of matrices. In our presentation, we are going to adopt the second method, for it is more straightforward, more elegant and much simpler than solving the Schrödinger equation. Unlike the examples seen up to now, solving the Schrödinger equation for the potential $V(x) = \frac{1}{2}m\omega x^2$ is no easy job. Before embarking on the second method, let us highlight the main steps involved in the first method.

Brief outline of the analytic method

This approach consists in using the power series method to solve the following differential (Schrödinger) equation:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + \frac{1}{2}m\omega^2x^2\psi(x) = E\psi(x), \tag{4.124}$$

which can be reduced to

$$\frac{d^2\psi(x)}{dx^2} + \left(\frac{2mE}{\hbar^2} - \frac{x^2}{x_0^4}\right)\psi(x) = 0, \tag{4.125}$$

where $x_0 = \sqrt{\hbar/(m\omega)}$ is a constant that has the dimensions of length; it sets the length scale of the oscillator, as will be seen later. The solutions of differential equations like (4.125) have been worked out by our mathematician colleagues well before the arrival of quantum mechanics (the solutions are expressed in terms of some special functions, the Hermite polynomials). The occurrence of the term $x^2\psi(x)$ in (4.125) suggests trying a Gaussian type solution[3]: $\psi(x) = f(x)\exp(-x^2/2x_0^2)$, where $f(x)$ is a function of x. Inserting this trial function into (4.125), we obtain a differential equation for $f(x)$. This new differential equation can be solved by expanding $f(x)$ out in a power series (i.e., $f(x) = \sum_{n=0}^{\infty} a_n x^n$, where a_n is just a coefficient), which when inserted into the differential equation leads to a recursion relation. By demanding the power series of $f(x)$ to terminate at some finite value of n (because the wave function $\psi(x)$ has to be finite everywhere, notably when $x \longrightarrow \pm\infty$), the recursion relation yields an expression for the energy eigenvalues which are discrete or quantized:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \qquad (n = 0, 1, 2, 3, \ldots). \tag{4.126}$$

After some calculations, we can show that the wave functions that are physically acceptable and that satisfy (4.125) are given by

$$\psi_n(x) = \frac{1}{\sqrt{\sqrt{\pi}2^n n! x_0}} e^{-x^2/2x_0^2} H_n\left(\frac{x}{x_0}\right), \tag{4.127}$$

where $H_n(y)$ are nth order polynomials called *Hermite polynomials*:

$$H_n(y) = (-1)^n e^{y^2}\frac{d^n}{dy^n}e^{-y^2}. \tag{4.128}$$

[3]Solutions of the type $\psi(x) = f(x)\exp(x^2/2x_0^2)$ are physically unacceptable, for they diverge when $x \longrightarrow \pm\infty$.

From this relation it is easy to calculate the first few polynomials:

$$\begin{aligned} H_0(y) &= 1, & H_1(y) &= 2y, \\ H_2(y) &= 4y^2 - 2, & H_3(y) &= 8y^3 - 12y, \\ H_4(y) &= 16y^4 - 48y^2 + 12, & H_5(y) &= 32y^5 - 160y^3 + 120y. \end{aligned} \tag{4.129}$$

We will deal with the physical interpretations of the harmonic oscillator results next when we study the second method – the Algebraic Method.

The Algebraic Method

Let us now show how to solve the harmonic oscillator eigenvalue problem using the algebraic method. For this, we need to rewrite the Hamiltonian (4.123) in terms of the two Hermitian, dimensionless operators $\hat{p} = \hat{P}/\sqrt{m\hbar\omega}$ and $\hat{q} = \hat{X}\sqrt{m\omega/\hbar}$:

$$\hat{H} = \frac{\hbar\omega}{2}(\hat{p}^2 + \hat{q}^2), \tag{4.130}$$

and then introduce two non-Hermitian, dimensionless operators:

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{q} + i\hat{p}), \qquad \hat{a}^\dagger = \frac{1}{\sqrt{2}}(\hat{q} - i\hat{p}). \tag{4.131}$$

The physical meaning of the operators $\hat{a}$ and $\hat{a}^\dagger$ will be examined later. Note that

$$\hat{a}^\dagger\hat{a} = \frac{1}{2}(\hat{q} - i\hat{p})(\hat{q} + i\hat{p}) = \frac{1}{2}(\hat{q}^2 + \hat{p}^2 + i\hat{q}\hat{p} - i\hat{p}\hat{q}) = \frac{1}{2}(\hat{q}^2 + \hat{p}^2) + \frac{i}{2}[\hat{q}, \hat{p}], \tag{4.132}$$

where, using $[\hat{X}, \hat{P}] = i\hbar$, we can verify that the commutator between $\hat{q}$ and $\hat{p}$ is

$$[\hat{q}, \hat{p}] = \left[\sqrt{\frac{m\omega}{\hbar}}\hat{X}, \frac{1}{\sqrt{\hbar m\omega}}\hat{P}\right] = \frac{1}{\hbar}\left[\hat{X}, \hat{P}\right] = i; \tag{4.133}$$

hence

$$\hat{a}^\dagger\hat{a} = \frac{1}{2}(\hat{q}^2 + \hat{p}^2) - \frac{1}{2} \tag{4.134}$$

or

$$\frac{1}{2}(\hat{q}^2 + \hat{p}^2) = \hat{a}^\dagger\hat{a} + \frac{1}{2}. \tag{4.135}$$

Inserting (4.135) into (4.130) we obtain

$$\boxed{\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) = \hbar\omega\left(\hat{N} + \frac{1}{2}\right) \qquad \text{with} \quad \hat{N} = \hat{a}^\dagger\hat{a},} \tag{4.136}$$

where $\hat{N}$ is known as the *number* operator or *occupation number* operator, which is clearly Hermitian.

Let us now derive the commutator $[\hat{a}, \hat{a}^\dagger]$. Since $[\hat{X}, \hat{P}] = i\hbar$ we have $[\hat{q}, \hat{p}] = \frac{1}{\hbar}[\hat{X}, \hat{P}] = i$; hence

$$[\hat{a}, \hat{a}^\dagger] = \frac{1}{2}[\hat{q} + i\hat{p}, \hat{q} - i\hat{p}] = -i[\hat{q}, \hat{p}] = 1 \tag{4.137}$$

or

$$\boxed{[\hat{a}, \hat{a}^\dagger] = 1.} \tag{4.138}$$

4.8.1 Energy Eigenvalues

Note that $\hat{H}$ as given by (4.136) commutes with $\hat{N}$, since $\hat{H}$ is linear in $\hat{N}$. Thus, $\hat{H}$ and $\hat{N}$ can have a set of joint eigenstates, to be denoted by $|\, n\rangle$:

$$\hat{N}\,|\, n\rangle = n\,|\, n\rangle \tag{4.139}$$

and

$$\hat{H}\,|\, n\rangle = E_n\,|\, n\rangle; \tag{4.140}$$

the states $|\, n\rangle$ are called energy eigenstates. Combining (4.136) and (4.140), we obtain the energy eigenvalues at once:

$$\boxed{E_n = \left(n + \frac{1}{2}\right)\hbar\omega.} \tag{4.141}$$

We will show later that n is a *positive integer*; it cannot have negative values.

The physical meaning of the operators $\hat{a}$, $\hat{a}^\dagger$, and $\hat{N}$ can now be clarified. First, we need the following two commutators that can be extracted from (4.138) and (4.136):

$$\boxed{[\hat{a}, \hat{H}] = \hbar\omega\hat{a}, \qquad [\hat{a}^\dagger, \hat{H}] = -\hbar\omega\hat{a}^\dagger.} \tag{4.142}$$

These commutation relations along with (4.140) lead to

$$\hat{H}\,(\hat{a}\,|\, n\rangle) \;\; = (\hat{a}\hat{H} - \hbar\omega\hat{a})\,|\, n\rangle \;\; = (E_n - \hbar\omega)\,(\hat{a}\,|\, n\rangle), \tag{4.143}$$

$$\hat{H}\left(\hat{a}^\dagger\,|\, n\rangle\right) = (\hat{a}^\dagger\hat{H} + \hbar\omega\hat{a}^\dagger)\,|\, n\rangle = (E_n + \hbar\omega)\,(\hat{a}^\dagger\,|\, n\rangle). \tag{4.144}$$

Thus, $\hat{a}\,|\, n\rangle$ and $\hat{a}^\dagger\,|\, n\rangle$ are eigenstates of $\hat{H}$ with eigenvalues $(E_n - \hbar\omega)$ and $(E_n + \hbar\omega)$, respectively. So the actions of $\hat{a}$ and $\hat{a}^\dagger$ on $|\, n\rangle$ generate new energy states that are lower and higher by one unit of $\hbar\omega$, respectively. As a result, $\hat{a}$ and $\hat{a}^\dagger$ are respectively known as the *lowering* and *raising* operators, or the *annihilation* and *creation* operators; they are also known as the *ladder* operators.

Let us now find out how the operators $\hat{a}$ and $\hat{a}^\dagger$ act on the energy eigenstates $|\, n\rangle$. Since $\hat{a}$ and $\hat{a}^\dagger$ do not commute with $\hat{N}$, the states $|\, n\rangle$ are eigenstates neither to $\hat{a}$ nor to $\hat{a}^\dagger$. Using (4.138) along with $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}$, we can show that

$$\boxed{[\hat{N}, \hat{a}] = -\hat{a}, \qquad [\hat{N}, \hat{a}^\dagger] = \hat{a}^\dagger;} \tag{4.145}$$

hence $\hat{N}\hat{a} = \hat{a}(\hat{N} - 1)$ and $\hat{N}\hat{a}^\dagger = \hat{a}^\dagger(\hat{N} + 1)$. Combining these relations with (4.139), we obtain

$$\hat{N}\,(\hat{a}\,|\, n\rangle) \;\; = \hat{a}(\hat{N} - 1)\,|\, n\rangle \;\; = (n - 1)\,(\hat{a}\,|\, n\rangle), \tag{4.146}$$

$$\hat{N}\left(\hat{a}^\dagger\,|\, n\rangle\right) = \hat{a}^\dagger(\hat{N} + 1)\,|\, n\rangle = (n + 1)(\hat{a}^\dagger\,|\, n\rangle). \tag{4.147}$$

These relations reveal that $\hat{a}\,|\, n\rangle$ and $\hat{a}^\dagger\,|\, n\rangle$ are eigenstates of $\hat{N}$ with eigenvalues $(n - 1)$ and $(n + 1)$, respectively. This implies that when $\hat{a}$ and $\hat{a}^\dagger$ operate on $|\, n\rangle$, respectively, they

decrease and increase n by one unit. That is, while the action of $\hat{a}$ on $|\, n\rangle$ generates a new state $|\, n-1\rangle$, the action of $\hat{a}^\dagger$ on $|\, n\rangle$ generates $|\, n+1\rangle$:

$$\hat{a}\,|\, n\rangle \propto |\, n-1\rangle, \qquad \hat{a}^\dagger\,|\, n\rangle \propto |\, n+1\rangle. \tag{4.148}$$

A combination of equations (4.146) and (4.147) with (4.148) lead, respectively, to

$$\hat{a}\,|\, n\rangle = b_n\,|\, n-1\rangle, \tag{4.149}$$

$$\hat{a}^\dagger\,|\, n\rangle = c_n\,|\, n+1\rangle. \tag{4.150}$$

where b_n and c_n are a constants to be determined from the requirement that the states $|\, n\rangle$ be normalized for all values of n.

First, to find b_n, it is easy to see that equation (4.149) leads to

$$\left(\langle n\,|\,\hat{a}^\dagger\right)\cdot(\hat{a}\,|\, n\rangle) = (\langle n-1\,|\, b_n^*)\cdot(b_n\,|\, n\rangle) = |b_n|^2\,\langle n-1\,|\, n-1\rangle = |b_n|^2 \tag{4.151}$$

and since $\hat{N}\,|\, n\rangle = \hat{a}^\dagger\hat{a}\,|\, n\rangle = n\,|\, n\rangle$ (see (4.139)), we have

$$\left(\langle n\,|\,\hat{a}^\dagger\right)\cdot(\hat{a}\,|\, n\rangle) = \langle n\,|\,\hat{a}^\dagger\hat{a}\,|\, n\rangle \equiv \langle n\,|\,\hat{N}\,|\, n\rangle = n\langle n\,|\, n\rangle = n. \tag{4.152}$$

When combined together, the last two equations yield

$$|b_n|^2 = n. \tag{4.153}$$

This implies that n, which is equal to the norm of $\hat{a}\,|\, n\rangle$ (see (4.152)), *cannot be negative*, $n \geq 0$, since the norm is a positive quantity; hence $b_n = \sqrt{n}$. Substituting $b_n = \sqrt{n}$ into (4.149), we end up with

$$\boxed{\hat{a}\,|\, n\rangle = \sqrt{n}\,|\, n-1\rangle.} \tag{4.154}$$

This equation shows that repeated applications of the operator $\hat{a}$ on $|\, n\rangle$ generate a sequence of eigenvectors $|\, n-1\rangle$, $|\, n-2\rangle$, $|\, n-3\rangle$, ... Since $n \geq 0$ and since $\hat{a}\,|\, 0\rangle = 0$, this sequence has to terminate at $n = 0$; this is true if we start with an integer value of n. But if we start with a noninteger n, the sequence will not terminate; hence it leads to eigenvectors with negative values of n. But as shown above, since n cannot be negative, we conclude that n has to be a *nonnegative integer*.

Second, to find c_n, let us invoke equation (4.150); it is easy to show that it leads to

$$(\langle n\,|\,\hat{a})\cdot\left(\hat{a}^\dagger\,|\, n\rangle\right) = (\langle n+1\,|\, c_n^*)\cdot(c_n\,|\, n+1\rangle) = |c_n|^2\,\langle n+1\,|\, n+1\rangle = |c_n|^2\,. \tag{4.155}$$

Now, from equation (4.138), we have $\hat{a}\hat{a}^\dagger = \hat{a}^\dagger\hat{a} + 1 \equiv N + 1$, and hence

$$\left(\langle n\,|\,\hat{a}^\dagger\right)\cdot(\hat{a}\,|\, n\rangle) = \langle n\,|\,\hat{a}\hat{a}^\dagger\,|\, n\rangle \equiv \langle n\,|\,\hat{N}+1\,|\, n\rangle = (n+1)\langle n+1\,|\, n+1\rangle = n+1. \tag{4.156}$$

Combining (4.155) and (4.156), we have

$$|c_n|^2 = n+1 \qquad \Longrightarrow \qquad c_n = \sqrt{n+1}. \tag{4.157}$$

Inserting (4.157) into (4.150), we obtain

$$\boxed{\hat{a}^\dagger \mid n\rangle = \sqrt{n+1}\mid n+1\rangle.} \tag{4.158}$$

This implies that repeated applications of $\hat{a}^\dagger$ on $\mid n\rangle$ generate an infinite sequence of eigenvectors $\mid n+1\rangle$, $\mid n+2\rangle$, $\mid n+3\rangle$, ... Since n is a positive integer, the energy spectrum of a harmonic oscillator as specified by (4.141) is therefore *discrete*:

$$\boxed{E_n = \left(n+\frac{1}{2}\right)\hbar\omega \qquad (n = 0,1,2,3,\ldots).} \tag{4.159}$$

This expression is similar to the one obtained from the first method (see Eq. (4.126)). The energy spectrum of the harmonic oscillator consists of energy levels that are equally spaced: $E_{n+1} - E_n = \hbar\omega$. This is Planck's famous equidistant energy idea—the energy of the radiation emitted by the oscillating charges (from the inside walls of the cavity) must come only in bundles (quanta) that are integral multiples of $\hbar\omega$—which, as mentioned in Chapter 1, led to the birth of quantum mechanics.

As expected for bound states of one-dimensional potentials, the energy spectrum is both discrete and nondegenerate. Once again, as in the case of the infinite square well potential, we encounter the zero-point energy phenomenon: the lowest energy eigenvalue of the oscillator is not zero but is instead equal to $E_0 = \hbar\omega/2$. It is called the *zero-point energy* of the oscillator, for it corresponds to $n = 0$. The zero-point energy of bound state systems cannot be zero, otherwise it would violate the uncertainty principle. For the harmonic oscillator, for instance, the classical minimum energy corresponds to $x = 0$ and $p = 0$; there would be no oscillations in this case. This would imply that we know simultaneously and with absolute precision both the position and the momentum of the system. This would contradict the uncertainty principle.

4.8.2 Energy Eigenstates

The algebraic or operator method can also be used to determine the energy eigenvectors. First, using (4.158), we see that the various eigenvectors can be written in terms of the ground state $\mid 0\rangle$ as follows:

$$\mid 1\rangle = \hat{a}^\dagger \mid 0\rangle, \tag{4.160}$$

$$\mid 2\rangle = \frac{1}{\sqrt{2}}\hat{a}^\dagger \mid 1\rangle = \frac{1}{\sqrt{2!}}\left(\hat{a}^\dagger\right)^2 \mid 0\rangle, \tag{4.161}$$

$$\mid 3\rangle = \frac{1}{\sqrt{3}}\hat{a}^\dagger \mid 2\rangle = \frac{1}{\sqrt{3!}}\left(\hat{a}^\dagger\right)^3 \mid 0\rangle, \tag{4.162}$$

$$\vdots$$

$$\mid n\rangle = \frac{1}{\sqrt{n}}\hat{a}^\dagger \mid n-1\rangle = \frac{1}{\sqrt{n!}}\left(\hat{a}^\dagger\right)^n \mid 0\rangle. \tag{4.163}$$

So, to find any excited eigenstate $\mid n\rangle$, we need simply to operate $\hat{a}^\dagger$ on $\mid 0\rangle$ n successive times.

Note that any set of kets $\mid n\rangle$ and $\mid n'\rangle$, corresponding to different eigenvalues, must be orthogonal, $\langle n' \mid n\rangle \sim \delta_{n',n}$, since $\hat{H}$ is Hermitian and none of its eigenstates is degenerate.

Moreover, the states $|0\rangle, |1\rangle, |2\rangle, |3\rangle, \ldots, |n\rangle, \ldots$ are *simultaneous eigenstates* of $\hat{H}$ and $\hat{N}$; the set $\{|n\rangle\}$ constitutes an orthonormal and complete basis:

$$\langle n' \mid n\rangle = \delta_{n',n}, \qquad \sum_{n=0}^{+\infty} |n\rangle\langle n| = 1. \tag{4.164}$$

4.8.3 Energy Eigenstates in Position Space

Let us now determine the harmonic oscillator wave function in the position representation.

Equations (4.160) to (4.163) show that, knowing the ground state wave function, we can determine any other eigenstate by successive applications of the operator $a^\dagger$ on the ground state. So let us first determine the ground state wave function in the position representation.

The operator $\hat{p}$, defined by $\hat{p} = \hat{P}/\sqrt{m\hbar\omega}$, is given in the position space by

$$\hat{p} = -\frac{i\hbar}{\sqrt{m\hbar\omega}}\frac{d}{dx} = -ix_0\frac{d}{dx}, \tag{4.165}$$

where, as mentioned above, $x_0 = \sqrt{\hbar/(m\omega)}$ is a constant that has the dimensions of length; it sets the length scale of the oscillator. We can easily show that the annihilation and creation operators $\hat{a}$ and $\hat{a}^\dagger$, defined in (4.131), can be written in the position representation as

$$\hat{a} = \frac{1}{\sqrt{2}}\left(\frac{\hat{X}}{x_0} + x_0\frac{d}{dx}\right) = \frac{1}{\sqrt{2}x_0}\left(\hat{X} + x_0^2\frac{d}{dx}\right), \tag{4.166}$$

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\left(\frac{\hat{X}}{x_0} - x_0\frac{d}{dx}\right) = \frac{1}{\sqrt{2}x_0}\left(\hat{X} - x_0^2\frac{d}{dx}\right). \tag{4.167}$$

Using (4.166) we can write the equation $\hat{a}\,|0\rangle = 0$ in the position space as

$$\langle x|\hat{a}\,|0\rangle = \frac{1}{\sqrt{2}x_0}\langle x|\hat{X} + x_0^2\frac{d}{dx}\,|0\rangle = \frac{1}{\sqrt{2}x_0}\left(x\psi_0(x) + x_0^2\frac{d\psi_0(x)}{dx}\right) = 0; \tag{4.168}$$

hence

$$\frac{d\psi_0(x)}{dx} = -\frac{x}{x_0^2}\psi_0(x), \tag{4.169}$$

where $\psi_0(x) = \langle x \mid 0\rangle$ represents the ground state wave function. The solution of this differential equation is

$$\psi_0(x) = A\exp\left(-\frac{x^2}{2x_0^2}\right), \tag{4.170}$$

where A is a constant that can be determined from the normalization condition

$$1 = \int_{-\infty}^{+\infty} dx\,|\psi_0(x)|^2 = A^2\int_{-\infty}^{+\infty} dx\exp\left(-\frac{x^2}{x_0^2}\right) = A^2\sqrt{\pi}x_0; \tag{4.171}$$

hence $A = (m\omega/(\pi\hbar))^{1/4} = 1/\sqrt{\sqrt{\pi}x_0}$. The normalized ground state wave function is then given by

$$\boxed{\psi_0(x) = \frac{1}{\sqrt{\sqrt{\pi}x_0}} \exp\left(-\frac{x^2}{2x_0^2}\right).} \tag{4.172}$$

This is a Gaussian function.

We can then obtain the wave function of any excited state by a series of applications of $\hat{a}^\dagger$ on the ground state. For instance, the first excited state is obtained by one single application of the operator $\hat{a}^\dagger$ of (4.167) on the ground state:

$$\begin{aligned} \langle x \mid 1\rangle &= \langle x|\hat{a}^\dagger \mid 0\rangle = \frac{1}{\sqrt{2}x_0}\left(x - x_0^2\frac{d}{dx}\right)\langle x \mid 0\rangle \\ &= \frac{1}{\sqrt{2}x_0}\left(x - x_0^2\left(-\frac{x}{x_0^2}\right)\right)\psi_0(x) = \frac{\sqrt{2}}{x_0}x\psi_0(x) \end{aligned} \tag{4.173}$$

or

$$\psi_1(x) = \frac{\sqrt{2}}{x_0}x\psi_0(x) = \sqrt{\frac{2}{\sqrt{\pi}x_0^3}}x\exp\left(-\frac{x^2}{2x_0^2}\right). \tag{4.174}$$

As for the eigenstates of the second and third excited states, we can obtain them by applying $\hat{a}^\dagger$ on the ground state twice and three times, respectively:

$$\langle x \mid 2\rangle = \frac{1}{\sqrt{2!}}\langle x|\left(a^\dagger\right)^2 \mid 0\rangle = \frac{1}{\sqrt{2!}}\left(\frac{1}{\sqrt{2}x_0}\right)^2\left(x - x_0^2\frac{d}{dx}\right)^2\psi_0(x), \tag{4.175}$$

$$\langle x \mid 3\rangle = \frac{1}{\sqrt{3!}}\langle x|\left(a^\dagger\right)^3 \mid 0\rangle = \frac{1}{\sqrt{3!}}\left(\frac{1}{\sqrt{2}x_0}\right)^3\left(x - x_0^2\frac{d}{dx}\right)^3\psi_0(x) \tag{4.176}$$

or

$$\psi_2(x) = \frac{1}{\sqrt{2\sqrt{\pi}x_0}}\left(\frac{2x^2}{x_0^2} - 1\right)\exp\left(-\frac{x^2}{2x_0^2}\right), \quad \psi_3(x) = \frac{1}{\sqrt{3\sqrt{\pi}x_0}}\left(\frac{2x^3}{x_0^3} - \frac{3x}{x_0}\right)\exp\left(-\frac{x^2}{2x_0^2}\right). \tag{4.177}$$

Similarly, using (4.163), (4.167), and (4.172), we can easily infer the energy eigenstate for the nth excited state:

$$\langle x \mid n\rangle = \frac{1}{\sqrt{n!}}\langle x|\left(a^\dagger\right)^n \mid 0\rangle = \frac{1}{\sqrt{n!}}\left(\frac{1}{\sqrt{2}x_0}\right)^n\left(x - x_0^2\frac{d}{dx}\right)^n\psi_0(x), \tag{4.178}$$

which in turn can be rewritten as

$$\boxed{\psi_n(x) = \frac{1}{\sqrt{\sqrt{\pi}2^n n!}}\frac{1}{x_0^{n+1/2}}\left(x - x_0^2\frac{d}{dx}\right)^n\exp\left(-\frac{x^2}{2x_0^2}\right).} \tag{4.179}$$

In summary, by successive applications of $\hat{a}^\dagger = (\hat{X} - x_0^2 d/dx)/(\sqrt{2}x_0)$ on $\psi_0(x)$, we can find the wave function of any excited state $\psi_n(x)$.

Oscillator wave functions and the Hermite polynomials

At this level, we can show that the wave function (4.179) derived from the algebraic method is similar to the one obtained from the first method (4.127). To see this, we simply need to use the following operator identity:

$$e^{-x^2/2}\left(x-\frac{d}{dx}\right)e^{x^2/2}=-\frac{d}{dx} \quad \text{or} \quad e^{-x^2/2x_0^2}\left(x-x_0^2\frac{d}{dx}\right)e^{x^2/2x_0^2}=-x_0^2\frac{d}{dx}. \tag{4.180}$$

An application of this operator n times leads at once to

$$e^{-x^2/2x_0^2}\left(x-x_0^2\frac{d}{dx}\right)^n e^{x^2/2x_0^2}=(-1)^n(x_0^2)^n\frac{d^n}{dx^n}, \tag{4.181}$$

which can be shown to yield

$$\left(x-x_0^2\frac{d}{dx}\right)^n e^{-x^2/2x_0^2}=(-1)^n(x_0^2)^n e^{x^2/2x_0^2}\frac{d^n}{dx^n}e^{-x^2/x_0^2}. \tag{4.182}$$

We can now rewrite the right-hand side of this equation as follows:

$$\begin{aligned}(-1)^n(x_0^2)^n e^{x^2/2x_0^2}\frac{d^n}{dx^n}e^{-x^2/x_0^2} &= x_0^n e^{-x^2/2x_0^2}\left[(-1)^n e^{x^2/x_0^2}\frac{d^n}{d(x/x_0)^n}e^{-x^2/x_0^2}\right]\\ &= x_0^n e^{-x^2/2x_0^2}\left[(-1)^n e^{y^2}\frac{d^n}{dy^n}e^{-y^2}\right]\\ &= x_0^n e^{-x^2/2x_0^2}H_n(y),\end{aligned} \tag{4.183}$$

where $y=x/x_0$ and where $H_n(y)$ are the Hermite polynomials listed in (4.128):

$$H_n(y)=(-1)^n e^{y^2}\frac{d^n}{dy^n}e^{-y^2}. \tag{4.184}$$

Note that the polynomials $H_{2n}(y)$ are even and $H_{2n+1}(y)$ are odd, since $H_n(-y)=(-1)^n H_n(y)$.

Inserting (4.183) into (4.182), we obtain

$$\left(x-x_0^2\frac{d}{dx}\right)^n e^{-x^2/2x_0^2}=x_0^n e^{-x^2/2x_0^2}H_n\left(\frac{x}{x_0}\right); \tag{4.185}$$

substituting this equation into (4.179), we can write the oscillator wave function in terms of the Hermite polynomials as follows:

$$\boxed{\psi_n(x)=\frac{1}{\sqrt{\sqrt{\pi}2^n n!x_0}}e^{-x^2/2x_0^2}H_n\left(\frac{x}{x_0}\right).} \tag{4.186}$$

This wave function is identical with the one obtained from the first method (see Eq. (4.127)).

Remark

This wave function is either even or odd depending on n; in fact, the functions $\psi_{2n}(x)$ are even (i.e., $\psi_{2n}(-x)=\psi_{2n}(x)$) and $\psi_{2n+1}(x)$ are odd (i.e., $\psi_{2n+1}(-x)=-\psi_{2n+1}(x)$) since, as can be inferred from Eq (4.129), the Hermite polynomials $H_{2n}(x)$ are even and $H_{2n+1}(x)$ are odd. This is expected because, as mentioned in Section 4.2.4, the wave functions of even one-dimensional potentials have definite parity. Figure 4.9 displays the shapes of the first few wave functions.

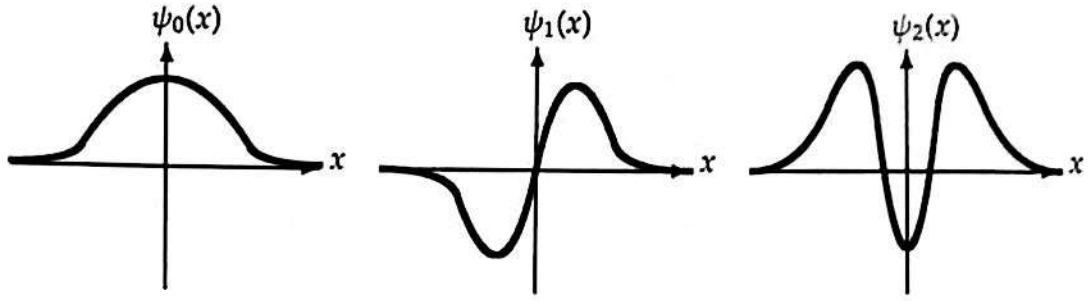


Figure 4.9: Shapes of the first three wave functions of the harmonic oscillator.

4.8.4 The Matrix Representation of Various Operators

Here we look at the matrix representation of several operators in the N-space. In particular, we focus on the representation of the operators $\hat{a}$, $\hat{a}^\dagger$, $\hat{X}$, and $\hat{P}$. First, since the states $|\, n\rangle$ are joint eigenstates of $\hat{H}$ and $\hat{N}$, it is easy to see from (4.139) and (4.141) that $\hat{H}$ and $\hat{N}$ are represented within the $\{|\, n\rangle\}$ basis by infinite diagonal matrices:

$$\langle n'|\hat{N}\,|\,n\rangle = n\delta_{n',n}, \qquad \langle n'|\hat{H}\,|\,n\rangle = \hbar\omega\left(n+\frac{1}{2}\right)\delta_{n',n}; \tag{4.187}$$

that is,

$$\hat{N} = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad \hat{H} = \frac{\hbar\omega}{2}\begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 3 & 0 & \cdots \\ 0 & 0 & 5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{4.188}$$

As for the operators $\hat{a}$, $\hat{a}^\dagger$, $\hat{X}$, $\hat{P}$, none of them are diagonal in the N-representation, since they do not commute with $\hat{N}$. The matrix elements of $\hat{a}$ and $\hat{a}^\dagger$ can be obtained from (4.154) and (4.158):

$$\langle n'|\hat{a}\,|\,n\rangle = \sqrt{n}\,\delta_{n',n-1}, \qquad \langle n'|\hat{a}^\dagger\,|\,n\rangle = \sqrt{n+1}\,\delta_{n',n+1}; \tag{4.189}$$

that is,

$$\hat{a} = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad \hat{a}^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ \sqrt{1} & 0 & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Now, let us find the N-representation of the position and momentum operators, $\hat{X}$ and $\hat{P}$. From (4.131) we can show that $\hat{X}$ and $\hat{P}$ are given in terms of $\hat{a}$ and $\hat{a}^\dagger$ as follows:

$$\hat{X} = \sqrt{\frac{\hbar}{2m\omega}}\left(\hat{a}+\hat{a}^\dagger\right), \quad \hat{P} = i\sqrt{\frac{m\hbar\omega}{2}}\left(\hat{a}^\dagger-\hat{a}\right). \tag{4.190}$$

Their matrix elements are given by

$$\langle n'|\hat{X}\,|\,n\rangle = \sqrt{\frac{\hbar}{2m\omega}}\left(\sqrt{n}\delta_{n',n-1} + \sqrt{n+1}\delta_{n',n+1}\right), \tag{4.191}$$

$$\langle n'|\hat{P}\,|\,n\rangle = i\sqrt{\frac{m\hbar\omega}{2}}\left(-\sqrt{n}\delta_{n',n-1} + \sqrt{n+1}\delta_{n',n+1}\right), \tag{4.192}$$

in particular

$$\langle n\,|\,\hat{X}\,|\,n\rangle = \langle n\,|\,\hat{P}\,|\,n\rangle = 0. \tag{4.193}$$

The matrices corresponding to $\hat{X}$ and $\hat{P}$ are thus given by

$$\hat{X} = \sqrt{\frac{\hbar}{2m\omega}}\begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \cdots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{4.194}$$

$$\hat{P} = i\sqrt{\frac{m\hbar\omega}{2}}\begin{pmatrix} 0 & -\sqrt{1} & 0 & 0 & \cdots \\ \sqrt{1} & 0 & -\sqrt{2} & 0 & \cdots \\ 0 & \sqrt{2} & 0 & -\sqrt{3} & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{4.195}$$

It is noteworthy to mention that last two matrices are infinite. Additionally, as demonstrated in Chapter 2, the momentum operator $\hat{P}$ is Hermitian, but not equal to its own complex conjugate; we can reach this result explicitly and easily from equation (4.195) which shows that $\hat{P}^\dagger = \hat{P}$ and $\hat{P}^* = -\hat{P}$. As for the position operator $\hat{X}$, however, it is both Hermitian and equal to its complex conjugate: from (4.194) we have that $\hat{X}^\dagger = \hat{X}^* = \hat{X}$.

Finally, we should mention that the eigenstates $|n\rangle$ are represented by infinite column matrices; the first few states can be written as

$$|\,0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \quad |\,1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \quad |\,2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \quad |\,3\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix}, \ldots . \tag{4.196}$$

The set of states $\{|\,n\rangle\}$ forms indeed a complete and orthonormal basis.

4.8.5 Expectation Values of Various Operators

Let us evaluate the expectation values for $\hat{X}^2$ and $\hat{P}^2$ in the N-representation:

$$\hat{X}^2 = \frac{\hbar}{2m\omega}\left(\hat{a}^2 + \hat{a}^{\dagger 2} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right) = \frac{\hbar}{2m\omega}\left(\hat{a}^2 + \hat{a}^{\dagger 2} + 2\hat{a}^\dagger\hat{a} + 1\right), \tag{4.197}$$

$$\hat{P}^2 = -\frac{m\hbar\omega}{2}\left(\hat{a}^2 + \hat{a}^{\dagger 2} - \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\right) = -\frac{m\hbar\omega}{2}\left(\hat{a}^2 + \hat{a}^{\dagger 2} - 2\hat{a}^\dagger\hat{a} - 1\right), \tag{4.198}$$

where we have used the fact that $\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} = 2\hat{a}^\dagger\hat{a} + 1$. Since the expectation values of $\hat{a}^2$ and $\hat{a}^{\dagger 2}$ are zero, $\langle n \mid \hat{a}^2 \mid n\rangle = \langle n \mid \hat{a}^{\dagger 2} \mid n\rangle = 0$, and $\langle n \mid \hat{a}^\dagger\hat{a} \mid n\rangle = n$, we have

$$\langle n \mid \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} \mid n\rangle = \langle n \mid 2\hat{a}^\dagger\hat{a} + 1 \mid n\rangle = 2n + 1; \tag{4.199}$$

hence

$$\langle n \mid \hat{X}^2 \mid n\rangle = \frac{\hbar}{2m\omega}\langle n \mid \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} \mid n\rangle = \frac{\hbar}{2m\omega}(2n+1), \tag{4.200}$$

$$\langle n \mid \hat{P}^2 \mid n\rangle = \frac{m\hbar\omega}{2}\langle n \mid \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} \mid n\rangle = \frac{m\hbar\omega}{2}(2n+1). \tag{4.201}$$

Comparing (4.200) and (4.201) we see that the expectation values of the potential and kinetic energies are equal and are also equal to half the total energy:

$$\boxed{\frac{m\omega^2}{2}\langle n \mid \hat{X}^2 \mid n\rangle = \frac{1}{2m}\langle n \mid \hat{P}^2 \mid n\rangle = \frac{1}{2}\langle n \mid \hat{H} \mid n\rangle.} \tag{4.202}$$

This result is known as the *Virial theorem.*

We can now easily calculate the product $\Delta x\Delta p$ from (4.200) and (4.201). Since $\langle\hat{X}\rangle = \langle\hat{P}\rangle = 0$ we have

$$\Delta x = \sqrt{\langle\hat{X}^2\rangle - \langle\hat{X}\rangle^2} = \sqrt{\langle\hat{X}^2\rangle} = \sqrt{\frac{\hbar}{2m\omega}(2n+1)}, \tag{4.203}$$

$$\Delta p = \sqrt{\langle\hat{P}^2\rangle - \langle\hat{P}\rangle^2} = \sqrt{\langle\hat{P}^2\rangle} = \sqrt{\frac{m\hbar\omega}{2}(2n+1)}; \tag{4.204}$$

hence

$$\Delta x\Delta p = \left(n + \frac{1}{2}\right)\hbar \quad\Longrightarrow\quad \Delta x\Delta p \geq \frac{\hbar}{2}, \tag{4.205}$$

since $n \geq 0$; this is the Heisenberg uncertainty principle.

4.9 Numerical Solution of the Schrödinger Equation

In this section we are going to show how to solve a one-dimensional Schrödinger equation numerically. The numerical solutions provide an idea about the properties of stationary states.

4.9.1 Numerical Procedure

We want to solve the following equation numerically:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi(x) = E\psi(x) \Longrightarrow \frac{d^2\psi}{dx^2} + k^2\psi(x) = 0, \tag{4.206}$$

where $k^2 = 2m[E - V(x)]/\hbar^2$.

First, divide the x-axis into a set of equidistant points with a spacing of $h_0 = \Delta x$, as shown in Figure 4.10a. The wave function $\psi(x)$ can be approximately described by its values at the

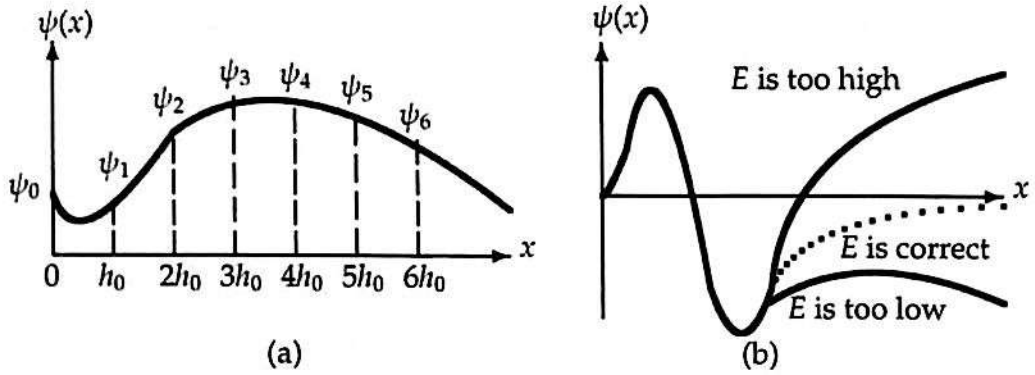


Figure 4.10: (a) Discretization of the wave function. (b) If the energy E used in the computation is too high (too low), the wave function will diverge as $x \to \pm\infty$; but at the appropriate value of E, the wave function converges to the correct values.

points of the grid (i.e., $\psi_0 = \psi(x = 0)$, $\psi_1 = \psi(h_0)$, $\psi_2 = \psi(2h_0)$, $\psi_3 = \psi(3h_0)$, and so on). The first derivative of ψ can then be approximated by

$$\frac{d\psi}{dx} \approx \frac{\psi_{n+1} - \psi_n}{h_0}. \tag{4.207}$$

An analogous approximation for the second derivative is actually a bit tricky. There are several methods to calculate it, but a very efficient procedure is called the *Numerov algorithm* (which is described in standard numerical analysis textbooks). In short, the second derivative is approximated by the so-called three-point difference formula:

$$\frac{\psi_{n+1} - 2\psi_n + \psi_{n-1}}{h_0^2} = \psi_n'' + \frac{h_0^2}{12}\psi_n'''' + 0(h_0^4). \tag{4.208}$$

From (4.206) we have

$$\psi_n'''' = \frac{d^2}{dx^2}(-k^2\psi)\bigg|_{x=x_n} = -\frac{(k^2\psi)_{n+1} - 2(k^2\psi)_n + (k^2\psi)_{n-1}}{h_0^2}. \tag{4.209}$$

Using $\psi_n'' = -k_n^2\psi_n$ and substituting (4.209) into (4.208) we can show that

$$\psi_{n+1} = \frac{2\left(1 - \frac{5}{12}h_0^2k_n^2\right)\psi_n - \left(1 + \frac{1}{12}h_0^2k_{n-1}^2\right)\psi_{n-1}}{1 + \frac{1}{12}h_0^2k_{n+1}^2}. \tag{4.210}$$

We can thus assign arbitrary values for ψ_0 and ψ_1; this is equivalent to providing the starting (or initial) values for $\psi(x)$ and $\psi'(x)$. Knowing ψ_0 and ψ_1, we can use (4.210) to calculate ψ_2, then ψ_3, then ψ_4, and so on. The solution of a linear equation, equation (4.210), for either ψ_{n+1} or ψ_{n-1} yields a recursion relation for integrating either *forward* or *backward* in x with a local error $O(h_0^6)$. In this way, the solution depends on two arbitrary constants, ψ_0 and ψ_1, as it should for any second-order differential equation (i.e., there are two linearly independent solutions).

The *boundary conditions* play a crucial role in solving any Schrödinger equation. Every boundary condition gives a linear homogeneous equation satisfied by the wave function or its derivative. For example, in the case of the infinite square well potential and the harmonic oscillator, the conditions $\psi(x_{min}) = 0$, $\psi(x_{max}) = 0$ are satisfied as follows:

- Infinite square well: $\psi(-a/2) = \psi(a/2) = 0$
- Harmonic oscillator: $\psi(-\infty) = \psi(+\infty) = 0$

4.9.2 Algorithm

To solve the Schrödinger equation with the boundary conditions $\psi(x_{min}) = \psi(x_{max}) = 0$, you may proceed as follows. Suppose you want to find the wave function, $\psi^{(n)}(x)$, and the energy E_n for the nth excited[4] state of a system:

- Take $\psi_0 = 0$ and choose ψ_1 (any small number you like), because the value of ψ_1 must be very close to that of ψ_0.
- Choose a *trial* energy E_n.
- With this value of the energy, E_n, together with ψ_0 and ψ_1, you can calculate iteratively the wave function at different values of x; that is, you can calculate $\psi_2, \psi_3, \psi_4, \ldots$. How? You need simply to inject $\psi_0 = 0$, ψ_1, and E_n into (4.210) and proceed incrementally to calculate ψ_2; then use ψ_1 and ψ_2 to calculate ψ_3; then use ψ_2 and ψ_3 to calculate ψ_4; and so on till you end up with the value of the wave function at $x_n = nh_0$, $\psi_n = \psi(nh_0)$.
- Next, you need to check whether the ψ_n you obtained is zero or not. If ψ_n is zero, this means that you have made the right choice for the trial energy. This value E_n can then be taken as a possible eigenenergy for the system; at this value of E_n, the wave function converges to the correct value (dotted curve in Figure 4.10b). Of course, it is highly unlikely to have chosen the correct energy from a first trial. In this case you need to proceed as follows. If the value of ψ_n obtained is a nonzero positive number or if it diverges, this means that the trial E_n you started with is larger than the correct eigenvalue (Figure 4.10b); on the other hand, if ψ_n is a negative nonzero number, this means that the E_n you started with is less than the true energy. If the ψ_n you end up with is a positive nonzero number, you need to start all over again with a smaller value of the energy. But if the ψ_n you end up with is negative, you need to start again with a larger value of E. You can continue in this way, improving every time, till you end up with a zero value for ψ_n. Note that in practice there is no way to get ψ_n exactly equal to zero. You may stop the procedure the moment ψ_n is sufficiently small; that is, you stop the iteration at the desired accuracy, say at 10^{-8} of its maximum value.

Example 4.3 (Numerical solution of the Schrödinger equation)
A proton is subject to a harmonic oscillator potential $V(x) = m\omega^2x^2/2$, $\omega = 5.34 \times 10^{21}s^{-1}$.

(a) Find the exact energies of the five lowest states (express them in MeV).

[4] We have denoted here the wave function of the nth excited state by $\psi^{(n)}(x)$ to distinguish it from the value of the wave function at $x_n = nh_0$, $\psi_n = \psi(nh_0)$.

Table 4.1: Exact and numerical energies for the five lowest states of the harmonic oscillator.

n	E_n^{Exact} (MeV)	$E_n^{Numeric}$ (MeV)
0.0	1.750 000	1.749 999 999 795
1.0	5.250 000	5.249 999 998 112
2.0	8.750 000	8.749 999 992 829
3.0	12.250 000	12.249 999 982 320
4.0	15.750 000	15.749 999 967 590

(b) Solve the Schrödinger equation *numerically* and find the energies of the five lowest states and compare them with the exact results obtained in (a). Note: You may use these quantities: rest mass energy of the proton $mc^2 \simeq 10^3$ MeV, $\hbar c \simeq 200$ MeV fm, and $\hbar\omega \simeq 3.5$ MeV.

Solution

(a) The exact energies can be calculated at once from $E_n = \hbar\omega(n + \frac{1}{2}) \simeq 3.5(n + \frac{1}{2})$ MeV. The results for the five lowest states are listed in Table 4.1.

(b) To obtain the numerical values, we need simply to make use of the Numerov relation (4.210), where $k_n^2(x) = 2m(E_n - \frac{1}{2}m\omega^2x^2)/\hbar^2$. The numerical values involved here can be calculated as follows:

$$\frac{m^2\omega^2}{\hbar^2} = \frac{(mc^2)^2(\hbar\omega)^2}{(\hbar c)^4} \simeq \frac{(10^3\,\text{MeV})^2(3.5\,\text{MeV})^2}{(200\,\text{MeV fm})^4} = 7.66 \times 10^{-4}\,\text{fm}^{-3}, \tag{4.211}$$

$$\frac{2m}{\hbar^2} = \frac{2mc^2}{(\hbar c)^2} \simeq \frac{2 \times 10^3\,\text{MeV}}{(200\,\text{MeV fm})^2} = 0.05\,\text{MeV}^{-1}\,\text{fm}^{-2}. \tag{4.212}$$

The boundary conditions for the harmonic oscillator imply that the wave function vanishes at $x = \pm\infty$, i.e., at $x_{min} = -\infty$ and $x_{max} = \infty$. How does one deal with infinities within a computer program? For this, we need to choose the numerical values of x_{min} and x_{max} in a way that the wave function would not feel the "edge" effects. That is, we simply need to assign numerical values to x_{min} and x_{max} so that they are far away from the turning points $x_{Left} = -\sqrt{2E_n/(m\omega^2)}$ and $x_{Right} = \sqrt{2E_n/(m\omega^2)}$, respectively. For instance, in the case of the ground state, where $E_0 = 1.75$ MeV, we have $x_{Left} = -3.38$ fm and $x_{Right} = 3.38$ fm; we may then take $x_{min} = -20$ fm and $x_{max} = 20$ fm. The wave function should be practically zero at $x = \pm 20$ fm.

To calculate the energies numerically for the five lowest states, a C++ computer code has been prepared (see Appendix C). The numerical results generated by this code are listed in Table 4.1; they are in excellent agreement with the exact results. Figure 4.11 displays the wave functions obtained from this code for the five lowest states for the proton moving in a harmonic oscillator potential (these plotted wave functions are normalized).

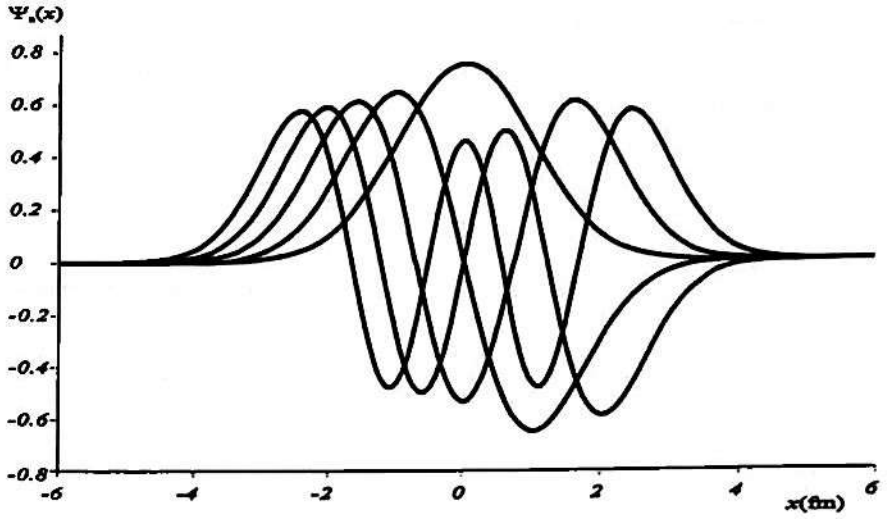


Figure 4.11: Wave functions $\psi_n(x)$ of the five lowest states of a harmonic oscillator potential in terms of x, where the x-axis values are in fm (obtained from the C++ code of Appendix C).

4.10 Solved Problems

Problem 4.1
A particle moving in one dimension is in a stationary state whose wave function

$$\psi(x) = \begin{cases} 0, & x < -a, \\ A(1 + \cos\frac{\pi x}{a}), & -a \le x \le a, \\ 0, & x > a, \end{cases}$$

where A and a are real constants.

(a) Is this a physically acceptable wave function? Explain.
(b) Find the magnitude of A so that $\psi(x)$ is normalized.
(c) Evaluate Δx and Δp. Verify that $\Delta x \Delta p \ge \hbar/2$.
(d) Find the classically allowed region.

Solution

(a) Since $\psi(x)$ is square integrable, single-valued, continuous, and has a continuous first derivative, it is indeed physically acceptable.

(b) Normalization of $\psi(x)$: using the relation $\cos^2 y = (1 + \cos 2y)/2$, we have

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} |\psi(x)|^2 dx = A^2 \int_{-a}^{a} dx \left[1 + 2\cos\frac{\pi x}{a} + \cos^2\left(\frac{\pi x}{a}\right)\right] \\ &= A^2 \int_{-a}^{a} dx \left[\frac{3}{2} + 2\cos\frac{\pi x}{a} + \frac{1}{2}\cos\frac{2\pi x}{a}\right] \\ &= \frac{3}{2}A^2 \int_{-a}^{a} dx = 3aA^2; \end{aligned} \tag{4.213}$$

hence $A = 1/\sqrt{3a}$.

(c) As $\psi(x)$ is even, we have $\langle\hat{X}\rangle = \int_{-a}^{+a} \psi^*(x)x\psi(x)\,dx = 0$, since the symmetric integral of an odd function (i.e., $\psi^*(x)x\psi(x)$ is odd) is zero. On the other hand, we also have $\langle\hat{P}\rangle = 0$ because $\psi(x)$ is real and even. We can thus write

$$\Delta x = \sqrt{\langle\hat{X}^2\rangle}, \qquad \Delta p = \sqrt{\langle\hat{P}^2\rangle}, \tag{4.214}$$

since $\Delta A = \sqrt{\langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2}$. The calculations of $\langle\hat{X}^2\rangle$ and $\langle\hat{P}^2\rangle$ are straightforward:

$$\begin{aligned}\langle\hat{X}^2\rangle &= \int_{-a}^{a} \psi^*(x)x^2\psi(x)\,dx = \frac{1}{3a}\int_{-a}^{a}\left[x^2 + 2x^2\cos\left(\frac{\pi x}{a}\right) + x^2\cos^2\left(\frac{\pi x}{a}\right)\right]dx \\ &= \frac{a^2}{6\pi^2}\left(2\pi^2 - 15\right),\end{aligned} \tag{4.215}$$

$$\begin{aligned}\langle\hat{P}^2\rangle &= -\hbar^2\int_{-a}^{+a}\psi(x)\frac{d^2\psi(x)}{dx^2}\,dx = \frac{\pi^2\hbar^2}{a^2}A^2\int_{-a}^{a}\left[\cos\frac{\pi x}{a} + \cos^2\left(\frac{\pi x}{a}\right)\right]dx \\ &= \frac{\pi^2\hbar^2}{3a^3}\int_{-a}^{a}\left[\frac{1}{2} + \cos\frac{\pi x}{a} + \frac{1}{2}\cos\frac{2\pi x}{a}\right]dx = \frac{\pi^2\hbar^2}{3a^2};\end{aligned} \tag{4.216}$$

hence $\Delta x = a\sqrt{1/3 - 5/(2\pi^2)}$ and $\Delta p = \pi\hbar/(\sqrt{3}a)$. We see that the uncertainties product

$$\Delta x\Delta p = \frac{\pi\hbar}{3}\sqrt{1 - \frac{15}{2\pi^2}} \tag{4.217}$$

satisfies Heisenberg's uncertainty principle, $\Delta x\Delta p > \hbar/2$.

(d) Since $d\psi^2/dx^2$ is zero at the inflection points, we have

$$\frac{d^2\psi}{dx^2} = -\frac{\pi^2}{a^2}A\cos\frac{\pi x}{a} = 0. \tag{4.218}$$

This relation holds when $x = \pm a/2$; hence the classically allowed region is defined by the interval between the inflection points $-a/2 \le x \le a/2$. That is, since $\psi(x)$ decays exponentially for $x > a/2$ and for $x < -a/2$, the energy of the system must be smaller than the potential. Classically, the system cannot be found in this region.

Problem 4.2

Consider a particle of mass m moving freely between $x = 0$ and $x = a$ inside an infinite square well potential.

(a) Calculate the expectation values $\langle\hat{X}\rangle_n$, $\langle\hat{P}\rangle_n$, $\langle\hat{X}^2\rangle_n$, and $\langle\hat{P}^2\rangle_n$, and compare them with their classical counterparts.

(b) Calculate the uncertainties product $\Delta x_n\Delta p_n$.

(c) Use the result of (b) to estimate the zero-point energy.

Solution

(a) Since $\psi_n(x) = \sqrt{2/a}\sin(n\pi x/a)$ and since it is a real function, we have $\langle\psi_n|\hat{P}|\psi_n\rangle = 0$ because for any real function $\phi(x)$ the integral $\langle\hat{P}\rangle = -i\hbar\int\phi^*(x)(d\phi(x)/dx)\,dx$ is imaginary and

this contradicts the fact that $\langle \hat{P} \rangle$ has to be real. On the other hand, the expectation values of $\hat{X}$, $\hat{X}^2$, and $\hat{P}^2$ are

$$\begin{aligned} \langle \psi_n | \hat{X} | \psi_n \rangle &= \int_0^a \psi_n^*(x) x \psi_n(x)\, dx = \frac{2}{a} \int_0^a x \sin^2\left(\frac{n\pi x}{a}\right) dx \\ &= \frac{1}{a} \int_0^a x \left[1 - \cos\left(\frac{2n\pi x}{a}\right)\right] dx = \frac{a}{2}, \end{aligned} \tag{4.219}$$

$$\begin{aligned} \langle \psi_n | \hat{X}^2 | \psi_n \rangle &= \frac{2}{a} \int_0^a x^2 \sin^2\left(\frac{n\pi x}{a}\right) dx = \frac{1}{a} \int_0^a x^2 \left[1 - \cos\left(\frac{2n\pi x}{a}\right)\right] dx \\ &= \frac{a^2}{3} - \frac{1}{a} \int_0^a x^2 \cos\left(\frac{2n\pi x}{a}\right) dx \\ &= \frac{a^2}{3} - \frac{1}{2n\pi} x^2 \sin\left(\frac{2n\pi x}{a}\right)\Big|_{x=0}^{x=a} + \frac{1}{n\pi} \int_0^a x \sin\left(\frac{2n\pi x}{a}\right) dx \\ &= \frac{a^2}{3} - \frac{a^2}{2n^2\pi^2}, \end{aligned} \tag{4.220}$$

$$\langle \psi_n | \hat{P}^2 | \psi_n \rangle = -\hbar^2 \int_0^a \psi_n^*(x) \frac{d^2 \psi_n(x)}{dx^2} dx = \frac{n^2\pi^2\hbar^2}{a^2} \int_0^a |\psi_n(x)|^2 dx = \frac{n^2\pi^2\hbar^2}{a^2}. \tag{4.221}$$

In deriving the previous three expressions, we have used integrations by parts. Since $E_n = n^2\pi^2\hbar^2/(2ma^2)$, we may write

$$\langle \psi_n | \hat{P}^2 | \psi_n \rangle = \frac{n^2\pi^2\hbar^2}{a^2} = 2mE_n. \tag{4.222}$$

To calculate the classical average values x_{av}, p_{av}, x_{av}^2, p_{av}^2, it is easy first to infer that $p_{av} = 0$ and $p_{av}^2 = 2mE$, since the particle moves to the right with *constant* momentum $p = mv$ and to the left with $p = -mv$. As the particle moves at constant speed, we have $x = vt$, hence

$$x_{av} = \frac{1}{T} \int_0^T x(t)\, dt = \frac{v}{T} \int_0^T t\, dt = v\frac{T}{2} = \frac{a}{2}, \tag{4.223}$$

$$x_{av}^2 = \frac{1}{T} \int_0^T x^2(t)\, dt = \frac{v^2}{T} \int_0^T t^2 dt = \frac{1}{3} v^2 T^2 = \frac{a^2}{3}, \tag{4.224}$$

where T is half[5] of the period of the motion, with $a = vT$.

We conclude that, while the average classical and quantum expressions for x, p and p^2 are identical, a comparison of (4.220) and (4.224) yields

$$\langle \psi_n | \hat{X}^2 | \psi_n \rangle = \frac{a^2}{3} - \frac{a^2}{2n^2\pi^2} = x_{av}^2 - \frac{a^2}{2n^2\pi^2}, \tag{4.225}$$

so that in the limit of large quantum numbers, the quantum expression $\langle \psi_n | \hat{X}^2 | \psi_n \rangle$ matches with its classical counterpart x_{av}^2: $\lim_{n\to\infty} \langle \psi_n | \hat{X}^2 | \psi_n \rangle = a^2/3 = x_{av}^2$.

[5] We may parameterize the other half of the motion by $x = -vt$, which when inserted in (4.223) and (4.224), where the variable t varies between $-T$ and 0, the integrals would yield the same results, namely $x_{av} = a/2$ and $x_{av}^2 = a^2/3$, respectively.

(b) The position and the momentum uncertainties can be calculated from (4.219) to (4.221):

$$\Delta x_n = \sqrt{\langle\psi_n|\hat{X}^2|\psi_n\rangle - \langle\psi_n|\hat{X}|\psi_n\rangle^2} = \sqrt{\frac{a^2}{3} - \frac{a^2}{2n^2\pi^2} - \frac{a^2}{4}} = a\sqrt{\frac{1}{12} - \frac{1}{2n^2\pi^2}}, \tag{4.226}$$

$$\Delta p_n = \sqrt{\langle\psi_n|\hat{P}^2|\psi_n\rangle - \langle\psi_n|\hat{P}|\psi_n\rangle^2} = \sqrt{\langle\psi_n|\hat{P}^2|\psi_n\rangle} = \frac{n\pi\hbar}{a}, \tag{4.227}$$

hence

$$\Delta x_n \Delta p_n = n\pi\hbar\sqrt{\frac{1}{12} - \frac{1}{2n^2\pi^2}}. \tag{4.228}$$

(c) Equation (4.227) shows that the momentum uncertainty for the ground state is not zero, but

$$\Delta p_1 = \frac{\pi\hbar}{a}. \tag{4.229}$$

This leads to a nonzero kinetic energy. Therefore, the lowest value of the particle's kinetic energy is of the order of $E_{min} \sim (\Delta p_1)^2/(2m) \sim \pi^2\hbar^2/(2ma^2)$. This value, which is in full agreement with the ground state energy, $E_1 = \pi^2\hbar^2/(2ma^2)$, is the zero-point energy of the particle.

Problem 4.3

An electron is moving freely inside a one-dimensional infinite potential box with walls at $x = 0$ and $x = a$. If the electron is initially in the ground state ($n = 1$) of the box and if we *suddenly* quadruple the size of the box (i.e., the right-hand side wall is moved instantaneously from $x = a$ to $x = 4a$), calculate the probability of finding the electron in:

(a) the ground state of the new box and

(b) the first excited state of the new box.

Solution

Initially, the electron is in the ground state of the box $x = 0$ and $x = a$; its energy and wave function are

$$E_1 = \frac{\pi^2\hbar^2}{2ma^2}, \qquad \phi_1(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{\pi x}{a}\right). \tag{4.230}$$

(a) Once in the new box, $x = 0$ and $x = 4a$, the ground state energy and wave function of the electron are

$$E_1' = \frac{\pi^2\hbar^2}{2m(4a)^2} = \frac{\pi^2\hbar^2}{32ma^2}, \qquad \psi_1(x) = \frac{1}{\sqrt{2a}}\sin\left(\frac{\pi x}{4a}\right). \tag{4.231}$$

The probability of finding the electron in $\psi_1(x)$ is

$$P(E_1') = |\langle\psi_1|\phi_1\rangle|^2 = \left|\int_0^a \psi_1^*(x)\phi_1(x)dx\right|^2 = \frac{1}{a^2}\left|\int_0^a \sin\left(\frac{\pi x}{4a}\right)\sin\left(\frac{\pi x}{a}\right)dx\right|^2; \tag{4.232}$$

the upper limit of the integral sign is a (and not $4a$) because $\phi_1(x)$ is limited to the region between 0 and a. Using the relation $\sin a \sin b = \frac{1}{2}\cos(a - b) - \frac{1}{2}\cos(a + b)$, we have

$\sin(\pi x/4a)\sin(\pi x/a) = \frac{1}{2}\cos(3\pi x/4a) - \frac{1}{2}\cos(5\pi x/4a)$; hence

$$P(E_1') = \frac{1}{a^2}\left|\frac{1}{2}\int_0^a \cos\left(\frac{3\pi x}{4a}\right)dx - \frac{1}{2}\int_0^a \cos\left(\frac{5\pi x}{4a}\right)dx\right|^2 = \frac{128}{15^2\pi^2} = 0.058 = 5.8\%. \quad (4.233)$$

(b) If the electron is in the first excited state of the new box, its energy and wave function are

$$E_2' = \frac{\pi^2\hbar^2}{8ma^2}, \qquad \psi_2(x) = \frac{1}{\sqrt{2a}}\sin\left(\frac{\pi x}{2a}\right). \quad (4.234)$$

The corresponding probability is

$$P(E_2') = |\langle\psi_2|\phi_1\rangle|^2 = \left|\int_0^a \psi_2^*(x)\phi_1(x)dx\right|^2 = \frac{1}{a^2}\left|\int_0^a \sin\left(\frac{\pi x}{2a}\right)\sin\left(\frac{\pi x}{a}\right)dx\right|^2 = \frac{16}{9\pi^2} = 0.18 = 18\%. \quad (4.235)$$

Problem 4.4

Consider a particle of mass m subject to an attractive delta potential $V(x) = -V_0\,\delta(x)$, where $V_0 > 0$ (V_0 has the dimensions of Energy×Distance).

(a) In the case of negative energies, show that this particle has only one bound state; find the binding energy and the wave function.

(b) Calculate the probability of finding the particle in the interval $-a \le x \le a$.

(c) What is the probability that the particle remains bound when V_0 is (i) halved suddenly, (ii) quadrupled suddenly?

(d) Study the scattering case (i.e., $E > 0$) and calculate the reflection and transmission coefficients as a function of the wave number k.

Solution

(a) Let us consider first the bound state case $E < 0$. We can write the Schrödinger equation as follows:

$$\frac{d^2\psi(x)}{dx^2} + \frac{2mV_0}{\hbar^2}\delta(x)\psi(x) + \frac{2mE}{\hbar^2}\psi(x) = 0. \quad (4.236)$$

Since $\delta(x)$ vanishes for $x \neq 0$, this equation becomes

$$\frac{d^2\psi(x)}{dx^2} + \frac{2mE}{\hbar^2}\psi(x) = 0. \quad (4.237)$$

The bound solutions require that $\psi(x)$ vanishes at $x = \pm\infty$; these bound solutions are given by

$$\psi(x) = \begin{cases} \psi_-(x) = Ae^{kx}, & x < 0, \\ \psi_+(x) = Be^{-kx}, & x > 0, \end{cases} \quad (4.238)$$

where $k = \sqrt{2m|E|}/\hbar$. Since $\psi(x)$ is continuous at $x = 0$, $\psi_-(0) = \psi_+(0)$, we have $A = B$. Thus, the wave function is given by $\psi(x) = Ae^{-k|x|}$; note that $\psi(x)$ is even.

The energy can be obtained from the discontinuity condition of the first derivative of the wave function, which in turn can be obtained by integrating (4.236) from $-\varepsilon$ to $+\varepsilon$,

$$\int_{-\varepsilon}^{+\varepsilon} dx \frac{d^2\psi(x)}{dx^2} + \frac{2mV_0}{\hbar^2}\int_{-\varepsilon}^{+\varepsilon} \delta(x)\psi(x)dx + \frac{2mE}{\hbar^2}\int_{-\varepsilon}^{+\varepsilon} \psi(x)dx = 0, \tag{4.239}$$

and then letting $\varepsilon \to 0$. Using the facts that

$$\int_{-\varepsilon}^{+\varepsilon} dx \frac{d^2\psi(x)}{dx^2} = \frac{d\psi(x)}{dx}\bigg|_{x=+\varepsilon} - \frac{d\psi(x)}{dx}\bigg|_{x=-\varepsilon} = \frac{d\psi_+(x)}{dx}\bigg|_{x=+\varepsilon} - \frac{d\psi_-(x)}{dx}\bigg|_{x=-\varepsilon} \tag{4.240}$$

and that $\int_{-\varepsilon}^{+\varepsilon} \psi(x)dx = 0$ (because $\psi(x)$ is even), we can rewrite (4.239) as follows:

$$\lim_{\varepsilon\to 0}\left(\frac{d\psi_+(x)}{dx}\bigg|_{x=+\varepsilon} - \frac{d\psi_-(x)}{dx}\bigg|_{x=-\varepsilon}\right)_{\varepsilon=0} + \frac{2mV_0}{\hbar^2}\psi(0) = 0, \tag{4.241}$$

since the wave function is continuous at $x = 0$, but its first derivative is not. Substituting (4.238) into (4.241) and using $A = B$, we obtain

$$(-2kA) + \frac{2mV_0}{\hbar^2}A = 0 \tag{4.242}$$

or $k = mV_0/\hbar^2$. But since $k = \sqrt{2m|E|/\hbar^2}$, we have $mV_0/\hbar^2 = \sqrt{2m|E|/\hbar^2}$, and since the energy is negative, we conclude that $E = -mV_0^2/(2\hbar^2)$. There is, therefore, only one bound state solution. As for the excited states, all of them are unbound. We may normalize $\psi(x)$,

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} \psi^*(x)\psi(x)\,dx = A^2\int_{-\infty}^{0} \exp(2kx)\,dx + A^2\int_0^{\infty} \exp(-2kx)\,dx \\ &= 2A^2\int_0^{\infty} \exp(-2kx)\,dx = \frac{A^2}{k}, \end{aligned} \tag{4.243}$$

hence $A = \sqrt{k}$. The normalized wave function is thus given by $\psi(x) = \sqrt{k}e^{-k|x|}$. So the energy and normalized wave function of the bound state are given by

$$E = -\frac{mV_0^2}{2\hbar^2}, \qquad \psi(x) = \sqrt{\frac{mV_0}{\hbar^2}}\exp\left(-\frac{mV_0}{\hbar^2}|x|\right). \tag{4.244}$$

(b) Since the wave function $\psi(x) = \sqrt{k}e^{-k|x|}$ is normalized, the probability of finding the particle in the interval $-a \le x \le a$ is given by

$$\begin{aligned} P &= \frac{\int_{-a}^{a}|\psi(x)|^2\,dx}{\int_{-\infty}^{\infty}|\psi(x)|^2\,dx} = \int_{-a}^{a}|\psi(x)|^2\,dx = k\int_{-a}^{a} e^{-2k|x|}dx \\ &= k\int_{-a}^{0} e^{2kx}dx + k\int_0^{a} e^{-2kx}dx = 2k\int_0^{a} e^{-2kx}dx \\ &= 1 - e^{-2ka} = 1 - e^{-2mV_0a/\hbar^2}. \end{aligned} \tag{4.245}$$

(c) If the strength of the potential changed suddenly from V_0 to V_1, the wave function will be given by $\psi_1(x) = \sqrt{mV_1/\hbar^2}\exp(-mV_1|x|/\hbar^2)$. The probability that the particle remains in the bound state $\psi_1(x)$ is

$$\begin{aligned} P &= |\langle\psi_1|\psi\rangle|^2 = \left|\int_{-\infty}^{\infty}\psi_1^*(x)\psi(x)\,dx\right|^2 \\ &= \left|\frac{m}{\hbar^2}\sqrt{V_0V_1}\int_{-\infty}^{\infty}\exp\left(-\frac{m(V_0+V_1)}{\hbar^2}|x|\right)dx\right|^2 \\ &= \left|2\frac{m}{\hbar^2}\sqrt{V_0V_1}\int_0^{\infty}\exp\left(-\frac{m(V_0+V_1)}{\hbar^2}x\right)dx\right|^2 = \frac{4V_0V_1}{(V_0+V_1)^2}. \end{aligned} \tag{4.246}$$

(i) In the case where the strength of the potential is halved, $V_1 = \frac{1}{2}V_0$, the probability that the particle remains bound is

$$P = \frac{2V_0^2}{(V_0+\frac{1}{2}V_0)^2} = \frac{8}{9} = 89\%. \tag{4.247}$$

(ii) When the strength is quadrupled, $V_1 = 4V_0$, the probability is given by

$$P = \frac{16V_0^2}{(5V_0)^2} = \frac{16}{25} = 64\%. \tag{4.248}$$

(d) The case $E > 0$ corresponds to a free motion and the energy levels represent a continuum. The solution of the Schrödinger equation for $E > 0$ is given by

$$\psi(x) = \begin{cases} \psi_-(x) = Ae^{ikx} + Be^{-ikx}, & x < 0, \\ \psi_+(x) = Ce^{ikx}, & x > 0, \end{cases} \tag{4.249}$$

where $k = \sqrt{2mE}/\hbar$; this corresponds to a plane wave incident from the left together with a reflected wave in the region $x < 0$, and only a transmitted wave for $x > 0$.

The values of the constants A and B are to be found from the continuity relations. On the one hand, the continuity of $\psi(x)$ at $x = 0$ yields

$$A + B = C \tag{4.250}$$

and, on the other hand, substituting (4.249) into (4.241), we end up with

$$ik(C - A + B) + \frac{2mV_0}{\hbar^2}C = 0. \tag{4.251}$$

Solving (4.250) and (4.251) for B/A and C/A, we find

$$\frac{B}{A} = \frac{-1}{1+\frac{ik\hbar^2}{mV_0}}, \qquad \frac{C}{A} = \frac{1}{1-\frac{imV_0}{\hbar^2 k}}. \tag{4.252}$$

Thus, the reflection and transmission coefficients are

$$R = \left|\frac{B}{A}\right|^2 = \frac{1}{1+\frac{\hbar^4k^2}{m^2V_0^2}} = \frac{1}{1+\frac{2\hbar^2E}{mV_0^2}}, \quad T = \left|\frac{C}{A}\right|^2 = \frac{1}{1+\frac{m^2V_0^2}{\hbar^4k^2}} = \frac{1}{1+\frac{mV_0^2}{2\hbar^2E}}, \tag{4.253}$$

with $R + T = 1$.

Problem 4.5

A particle of mass m is subject to an attractive double-delta potential $V(x) = -V_0\delta(x-a) - V_0\delta(x+a)$, where $V_0 > 0$. Consider only the case of negative energies.

(a) Obtain the wave functions of the bound states.

(b) Derive the eigenvalue equations.

(c) Specify the number of bound states and the limit on their energies. Is the ground state an even state or an odd state?

(d) Estimate the ground state energy for the limits $a \to 0$ and $a \to \infty$.

Solution

(a) The Schrödinger equation for this problem is

$$\frac{d^2\psi(x)}{dx^2} + \frac{2mV_0}{\hbar^2}\left[\delta(x-a) + \delta(x+a)\right]\psi(x) + \frac{2mE}{\hbar^2}\psi(x) = 0. \tag{4.254}$$

For $x \neq \pm a$ this equation becomes

$$\frac{d^2\psi(x)}{dx^2} + \frac{2mE}{\hbar^2}\psi(x) = 0 \qquad \text{or} \qquad \frac{d^2\psi(x)}{dx^2} - k^2\psi(x) = 0, \tag{4.255}$$

where $k^2 = -2mE/\hbar^2 = 2m|E|/\hbar^2$, since this problem deals only with the bound states $E < 0$.

Since the potential is symmetric, $V(-x) = V(x)$, the wave function is either even or odd; we will denote the even states by $\psi_+(x)$ and the odd states by $\psi_-(x)$. The bound state solutions for $E < 0$ require that $\psi_\pm(x)$ vanish at $x = \pm\infty$:

$$\psi_\pm(x) = \begin{cases} Ae^{-kx}, & x > a, \\ \frac{B}{2}\left(e^{kx} \pm e^{-kx}\right), & -a < x < a, \\ \pm Ae^{kx}, & x < -a; \end{cases} \tag{4.256}$$

hence

$$\psi_+(x) = \begin{cases} Ae^{-kx}, \\ B\cosh kx, \\ Ae^{kx}, \end{cases} \qquad \psi_-(x) = \begin{cases} Ae^{-kx}, & x > a, \\ B\sinh kx, & -a < x < a, \\ -Ae^{kx}, & x < -a. \end{cases} \tag{4.257}$$

The shapes of $\psi_\pm(x)$ are displayed in Figure 4.12.

(b) As for the energy eigenvalues, they can be obtained from the boundary conditions. The continuity condition at $x = a$ of $\psi_+(x)$ leads to

$$Ae^{-ka} = B\cosh ka \tag{4.258}$$

and that of $\psi_-(x)$ leads to

$$Ae^{-ka} = B\sinh ka. \tag{4.259}$$

To obtain the discontinuity condition for the first derivative of $\psi_+(x)$ at $x = a$, we need to integrate (4.254):

$$\lim_{\varepsilon\to 0}\left[\psi'_+(a+\varepsilon) - \psi'_+(a-\varepsilon)\right] + \frac{2mV_0}{\hbar^2}\psi_+(a) = 0; \tag{4.260}$$

hence

$$-kAe^{-ka} - kB\sinh ka + \frac{2mV_0}{\hbar^2}Ae^{-ka} = 0 \implies A\left(\frac{2mV_0}{k\hbar^2} - 1\right)e^{-ka} = B\sinh ka. \tag{4.261}$$

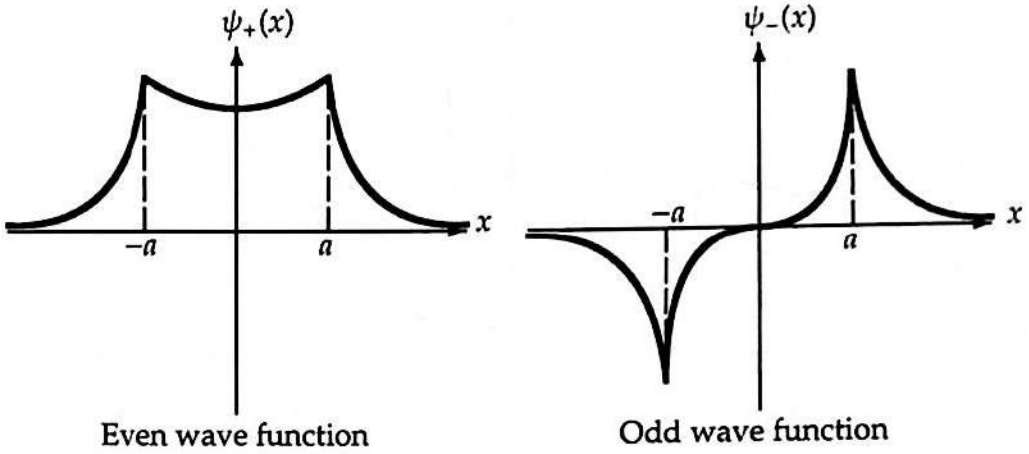


Figure 4.12: Shapes of the even and odd wave functions for $V(x) = -V_0\delta(x-a) - V_0\delta(x+a)$.

Similarly, the continuity of the first derivative of $\psi_-(x)$ at $x = a$ yields

$$-kAe^{-ka} - kB\cosh ka + \frac{2mV_0}{\hbar^2}Ae^{-ka} = 0 \implies A\left(\frac{2mV_0}{k\hbar^2} - 1\right)e^{-ka} = B\cosh ka. \tag{4.262}$$

Dividing (4.261) by (4.258) we obtain the eigenvalue equation for the even solutions:

$$\frac{2mV_0}{k\hbar^2} - 1 = \tanh ka \quad \implies \quad \tanh y = \frac{\gamma}{y} - 1, \tag{4.263}$$

where $y = ka$ and $\gamma = 2maV_0/\hbar^2$. The eigenvalue equation for the odd solutions can be obtained by dividing (4.262) by (4.259):

$$\frac{2mV_0}{k\hbar^2} - 1 = \coth ka \quad \implies \quad \coth y = \frac{\gamma}{y} - 1 \quad \implies \quad \tanh y = \left(\frac{\gamma}{y} - 1\right)^{-1}, \tag{4.264}$$

because $\coth y = 1/\tanh y$.

To obtain the energy eigenvalues for the even and odd solutions, we need to solve the transcendental equations (4.263) and (4.264). These equations can be solved graphically. In what follows, let us determine the upper and lower limits of the energy for both the even and odd solutions.

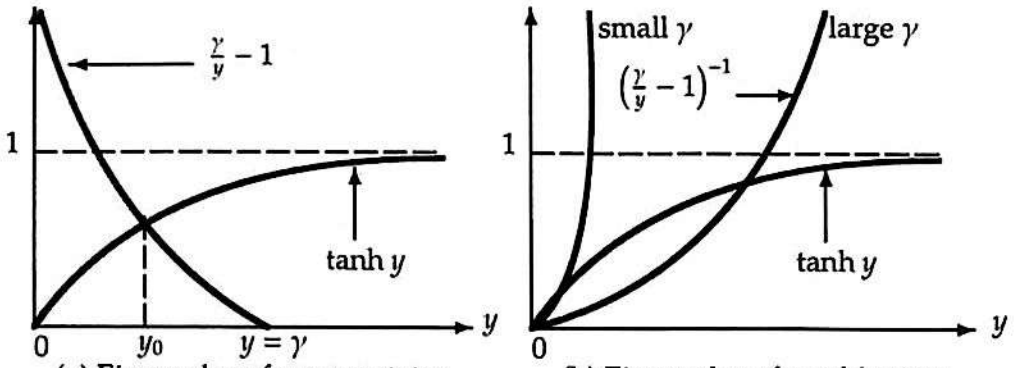
(c) To find the number of bound states and the limits on the energy, let us consider the even and odd states separately.

Energies corresponding to the even solutions

There is only one bound state, since the curves $\tanh y$ and $\gamma/y - 1$ intersect only once (Figure 4.13a); we call this point $y = y_0$. When $y = \gamma$ we have $\gamma/y - 1 = 0$, while $\tanh\gamma > 0$. Therefore $y_0 < \gamma$. On the other hand, since $\tanh y_0 < 1$ we have $\gamma/y_0 - 1 < 1$ or $y_0 > \gamma/2$. We conclude then that $\gamma/2 < y_0 < \gamma$ or

$$\frac{\gamma}{2} < y_0 < \gamma \quad \implies \quad -\frac{2mV_0^2}{\hbar^2} < E_{even} < -\frac{mV_0^2}{2\hbar^2}. \tag{4.265}$$

In deriving this relation, we have used the fact that $\gamma^2/4 < y_0^2 < \gamma^2$ where $\gamma = 2maV_0/\hbar^2$ and $y_0^2 = k_0^2a^2 = -2ma^2E_{even}/\hbar^2$. So there is always one even bound state, the ground state, whose energy lies within the range specified by (4.265).



(a) Eigenvalues for even states (b) Eigenvalues for odd states

Figure 4.13: Graphical solutions of the eigenvalue equations for the even states and the odd states for the double-delta potential $V(x) = -V_0\delta(x-a) - V_0\delta(x+a)$.

Energies corresponding to the odd solutions

As shown in Figure 4.13b, if the slope of $(\gamma/y-1)^{-1}$ at $y = 0$ is smaller than the slope of $\tanh y$, i.e.,

$$\left.\frac{d}{dy}\left(\frac{\gamma}{y}-1\right)^{-1}\right|_{y=0} < \left.\frac{d\tanh y}{dy}\right|_{y=0} \quad \Longrightarrow \quad \frac{1}{\gamma} < 1 \tag{4.266}$$

or

$$\gamma > 1 \quad \Longrightarrow \quad V_0 > \frac{\hbar^2}{2ma}, \tag{4.267}$$

there would be only one bound state because the curves $\tanh y$ and $(\gamma/y-1)^{-1}$ would intersect once. But if $\gamma < 1$ or $V_0 < \hbar^2/(2ma)$, there would be no odd bound states, for the curves of $\tanh y$ and $(\gamma/y-1)^{-1}$ would never intersect.

Note that if $y = \gamma/2$ we have $(\gamma/y-1)^{-1} = 1$. Thus the intersection of $\tanh y$ and $(\gamma/y-1)^{-1}$, if it takes place at all, has to take place for $y < \gamma/2$. That is, the odd bound states occur only when

$$y < \frac{\gamma}{2} \quad \Longrightarrow \quad E_{odd} > -\frac{mV_0^2}{2\hbar^2}. \tag{4.268}$$

A comparison of (4.265) and (4.268) shows that the energies corresponding to even states are smaller than those of odd states:

$$E_{even} < E_{odd}. \tag{4.269}$$

Thus, the even bound state is the ground state. Using this result, we may infer (a) if $\gamma < 1$ there are no odd bound states, but there is always one even bound state, the ground state; (b) if $\gamma > 1$ there are two bound states: the ground state (even) and the first excited state (odd).

We may summarize these results as follows:

If $\gamma < 1$ or $V_0 < \dfrac{\hbar^2}{2ma}$ $\longrightarrow$ there is only one bound state. (4.270)

If $\gamma > 1$ or $V_0 > \dfrac{\hbar^2}{2ma}$ $\longrightarrow$ there are two bound states. (4.271)

(d) In the limit $a \longrightarrow 0$ we have $y \longrightarrow 0$ and $\gamma \longrightarrow 0$; hence the even transcendental equation $\tanh y = \gamma/y - 1$ reduces to $y \simeq \gamma/y - 1$ or $y = \gamma$, which in turn leads to $y^2 = (ka)^2 = \gamma^2$ or $-2ma^2E_{even}/\hbar^2 = (2maV_0/\hbar^2)^2$:

$$E_{even} = -\frac{2mV_0^2}{\hbar^2}. \tag{4.272}$$

Note that in the limit $a \longrightarrow 0$, the potential $V(x) = -V_0\delta(x-a) - V_0\delta(x+a)$ reduces to $V(x) = -2V_0\delta(x)$. We can see that the ground state energy (4.244) of the single-delta potential is identical with (4.272) provided we replace V_0 in (4.244) by $2V_0$.

In the limit $a \longrightarrow \infty$, we have $y \longrightarrow \infty$ and $\gamma \longrightarrow \infty$; hence $\tanh y = \gamma/y - 1$ reduces to $1 \simeq \gamma/y - 1$ or $y = \gamma/2$. This leads to $y^2 = (ka)^2 = \gamma^2/4$ or $-2ma^2E_{even}/\hbar^2 = (maV_0/\hbar^2)^2$:

$$E_{even} = -\frac{mV_0^2}{2\hbar^2}. \tag{4.273}$$

This relation is identical with that of the single-delta potential (4.244).

Problem 4.6
Consider a particle of mass m subject to the potential

$$V(x) = \begin{cases} \infty, & x \le 0, \\ -V_0\delta(x-a), & x > 0, \end{cases}$$

where $V_0 > 0$. Discuss the existence of bound states in terms of the size of a.

Solution
The Schrödinger equation for $x > 0$ is

$$\frac{d^2\psi(x)}{dx^2} + \left[\frac{2mV_0}{\hbar^2}\delta(x-a) - k^2\right]\psi(x) = 0, \tag{4.274}$$

where $k^2 = -2mE/\hbar^2$, since we are looking here at the bound states only, $E < 0$. The solutions of this equation are

$$\psi(x) = \begin{cases} \psi_1(x) = Ae^{kx} + Be^{-kx}, & 0 < x < a, \\ \psi_2(x) = Ce^{-kx}, & x > a. \end{cases} \tag{4.275}$$

The energy eigenvalues can be obtained from the boundary conditions. As the wave function vanishes at $x = 0$, we have

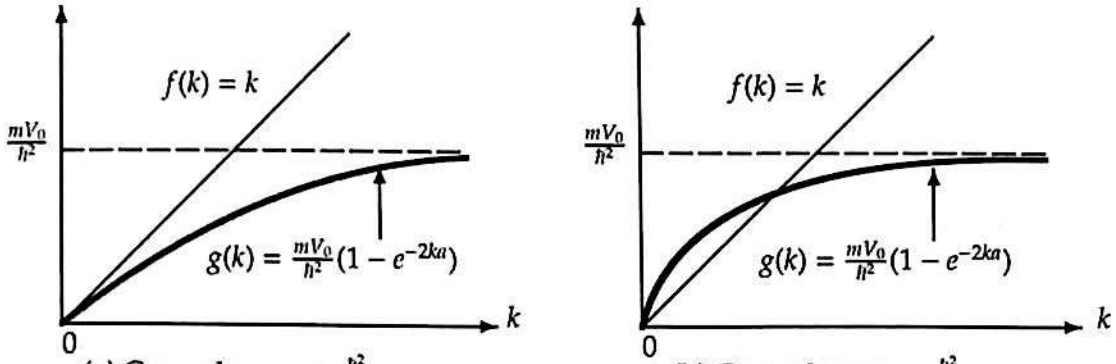
$$\psi_1(0) = 0 \quad \Longrightarrow \quad A + B = 0 \quad \Longrightarrow \quad B = -A. \tag{4.276}$$

The continuity condition at $x = a$ of $\psi(x)$, $\psi_1(a) = \psi_2(a)$, leads to

$$Ae^{ka} - Ae^{-ka} = Ce^{-ka}. \tag{4.277}$$

To obtain the discontinuity condition for the first derivative of $\psi(x)$ at $x = a$, we need to integrate (4.274):

$$\lim_{\varepsilon \to a}\left[\psi_2'(a+\varepsilon) - \psi_1'(a-\varepsilon)\right] + \frac{2mV_0}{\hbar^2}\psi_2(a) = 0 \tag{4.278}$$



(a) Case where $a < \frac{\hbar^2}{2mV_0}$ (b) Case where $a > \frac{\hbar^2}{2mV_0}$

Figure 4.14: Graphical solutions of $f(k) = g(k)$ or $k = (mV_0/\hbar^2)(1 - e^{-2ka})$. If the slope of $g(k)$ is smaller than 1, i.e., $a < \hbar^2/(2mV_0)$, no bound state will exist, but if the slope of $g(k)$ is greater than 1, i.e., $a > \frac{\hbar^2}{2mV_0}$, there will be only one bound state.

or

$$-kCe^{-ka} - kAe^{ka} - kAe^{-ka} + \frac{2mV_0}{\hbar^2}Ce^{-ka} = 0 \tag{4.279}$$

Substituting $Ce^{-ka} = Ae^{ka} - Ae^{-ka}$ or (4.277) into (4.279) we have

$$-kAe^{ka} + kAe^{-ka} - kAe^{ka} - kAe^{-ka} + \frac{2mV_0}{\hbar^2}\left(Ae^{ka} - Ae^{-ka}\right) = 0. \tag{4.280}$$

From this point on, we can proceed in two different, yet equivalent, ways. These two methods differ merely in the way we exploit (4.280). For completeness of the presentation, let us discuss both methods.

First method

The second and fourth terms of (4.280) cancel each other, so we can reduce it to

$$-kAe^{ka} - kAe^{ka} + \frac{2mV_0}{\hbar^2}\left(Ae^{ka} - Ae^{-ka}\right) = 0, \tag{4.281}$$

which in turn leads to the following transcendental equation:

$$k = \frac{mV_0}{\hbar^2}\left(1 - e^{-2ka}\right). \tag{4.282}$$

The energy eigenvalues are given by the intersection of the curves $f(k) = k$ and $g(k) = mV_0(1 - e^{-2ka})/\hbar^2$. As the slope of $f(k)$ is equal to 1, if the slope of $g(k)$ at $k = 0$ is smaller than 1 (i.e., $a < \hbar^2/(2mV_0)$), there will be no bound states (Figure 4.14a). But if the slope of $g(k)$ is greater than 1 (i.e., $a > \hbar^2/(2mV_0)$),

$$\left.\frac{dg(k)}{dk}\right|_{k=0} > 1 \qquad \text{or} \qquad a > \frac{\hbar^2}{2mV_0}, \tag{4.283}$$

and there will be one bound state (Figure 4.14b).

Second method

We simply combine the first and second terms of (4.280) to generate $-2kA\sinh(ka)$; the third

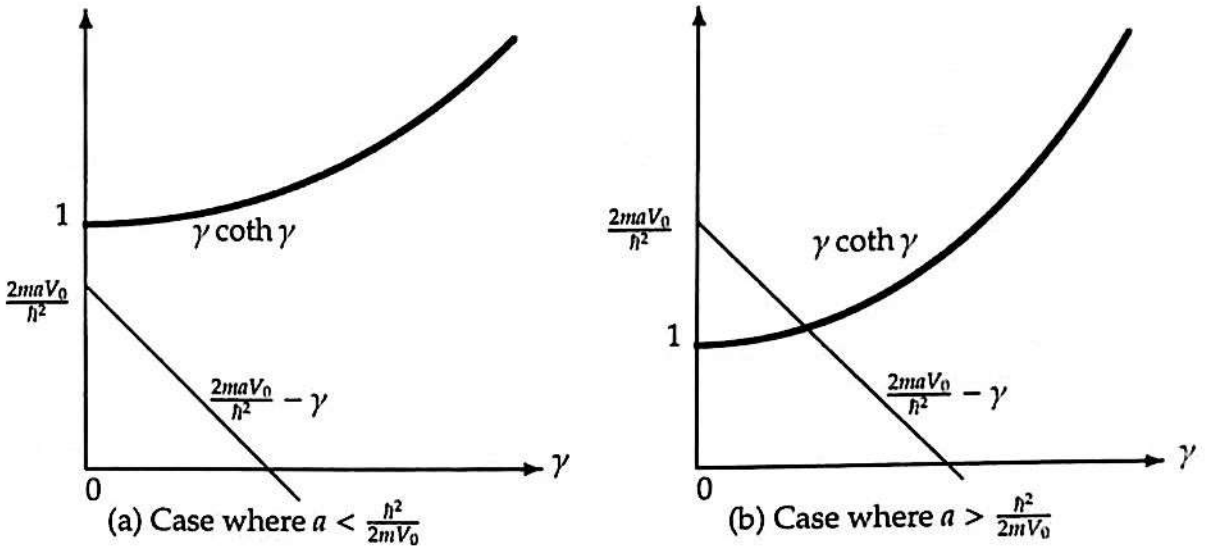


Figure 4.15: Graphical solutions of $h(\gamma) = u(\gamma)$, with $\gamma = ka$, $h(\gamma) = \gamma \coth \gamma$, and $u(\gamma) = 2mV_0a/\hbar^2 - \gamma$. If $a < 2mV_0/\hbar^2$ there is no bound state. If $a > 2mV_0/\hbar^2$ there is one bound state.

and fourth terms yield $-2kA\cosh(ka)$; and the fifth and sixth terms lead to $2A(2mV_0/\hbar^2)\sinh ka$. Hence

$$-2kA \sinh ka - 2kA \cosh ka + (2A)\frac{2mV_0}{\hbar^2} \sinh ka = 0, \tag{4.284}$$

which leads to

$$\gamma \coth \gamma = \frac{2mV_0}{\hbar^2}a - \gamma, \tag{4.285}$$

where $\gamma = ka$. The energy eigenvalues are given by the intersection of the curves $h(\gamma) = \gamma \coth \gamma$ and $u(\gamma) = 2mV_0a/\hbar^2 - \gamma$. As displayed in Figure 4.15a, if $a < 2mV_0/\hbar^2$, no bound state solution will exist, since the curves of $h(\gamma)$ and $u(\gamma)$ do not intersect. But if $a > 2mV_0/\hbar^2$, the curves intersect only once; hence there will be one bound state (Figure 4.15b).

We may summarize the results as follows:

$$a < \frac{\hbar^2}{2mV_0} \implies \text{no bound states,} \tag{4.286}$$

$$a > \frac{\hbar^2}{2mV_0} \implies \text{one bound state.} \tag{4.287}$$

Problem 4.7

A particle of mass m, besides being confined to move inside an infinite square well potential of size a with walls at $x = 0$ and $x = a$, is subject to a delta potential of strength V_0

$$V(x) = \begin{cases} V_0\delta(x - a/2), & 0 < x < a, \\ \infty, & \text{elsewhere,} \end{cases}$$

where $V_0 > 0$. Show how to calculate the energy levels of the system in terms of V_0 and a.

Solution
The Schrödinger equation

$$\frac{d^2\psi(x)}{dx^2} + \frac{2mV_0}{\hbar^2}\delta\left(x - \frac{a}{2}\right)\psi(x) + \frac{2mE}{\hbar^2}\psi(x) = 0 \tag{4.288}$$

can be written for $x \neq a/2$ as

$$\frac{d^2\psi(x)}{dx^2} + \frac{2mE}{\hbar^2}\psi(x) = 0. \tag{4.289}$$

The solutions of this equation must vanish at $x = 0$ and $x = a$:

$$\psi(x) = \begin{cases} \psi_L(x) = A\sin kx, & 0 \leq x < a/2, \\ \psi_R(x) = B\sin k(x-a), & a/2 < x \leq a, \end{cases} \tag{4.290}$$

where $k = \sqrt{2m|E|}/\hbar$. The continuity of $\psi(x)$ at $x = a/2$, $\psi_L(a/2) = \psi_R(a/2)$, leads to $A\sin(a/2) = -B\sin(a/2)$; hence $B = -A$. The wave function is thus given by

$$\psi(x) = \begin{cases} \psi_L(x) = A\sin kx, & 0 \leq x < a/2, \\ \psi_R(x) = -A\sin k(x-a), & a/2 < x \leq a, \end{cases} \tag{4.291}$$

The energy eigenvalues can be found from the discontinuity condition of the first derivative of the wave function, which in turn can be obtained by integrating (4.288) from $a/2 - \varepsilon$ to $a/2 + \varepsilon$ and then letting $\varepsilon \to 0$:

$$\lim_{\varepsilon\to 0}\left(\left.\frac{d\psi_R(x)}{dx}\right|_{x=a/2+\varepsilon} - \left.\frac{d\psi_L(x)}{dx}\right|_{x=a/2-\varepsilon}\right)_{\varepsilon=0} + \frac{2mV_0}{\hbar^2}\psi\left(\frac{a}{2}\right) = 0. \tag{4.292}$$

Substituting (4.291) into (4.292) we obtain

$$-kA\cos\left[k\left(\frac{a}{2} - a\right)\right] - kA\cos\left(k\frac{a}{2}\right) + A\frac{2mV_0}{\hbar^2}\sin\left(k\frac{a}{2}\right) = 0 \tag{4.293}$$

or

$$\tan\left(k\frac{a}{2}\right) = \frac{\hbar^2 k}{mV_0} \quad \Longrightarrow \quad \tan\left(\sqrt{\frac{ma^2|E|}{2\hbar^2}}\right) = \sqrt{\frac{2\hbar^2|E|}{mV_0^2}}. \tag{4.294}$$

This is a transcendental equation for the energy; its solutions, which can be obtained numerically or graphically, yield the values of E.

Problem 4.8
Using the uncertainty principle, show that the lowest energy of an oscillator is $\hbar\omega/2$.

Solution
The motion of the particle is confined to the region $-a/2 \leq x \leq a/2$; that is, $\Delta x \simeq a$. Then as a result of the uncertainty principle, the lowest value of this particle's momentum is $\hbar/(2\Delta x) \simeq \hbar/(2a)$. The total energy as a function of a is

$$E(a) \simeq \frac{1}{2m}\left(\frac{\hbar}{2a}\right)^2 + \frac{1}{2}m\omega^2 a^2. \tag{4.295}$$

The minimization of E with respect to a,

$$0 = \left.\frac{dE}{da}\right|_{a=a_0} = -\frac{\hbar^2}{4ma_0^3} + m\omega^2 a_0, \tag{4.296}$$

gives $a_0 = \sqrt{\hbar/2m\omega}$ and hence $E(a_0) \simeq \hbar\omega/2$; this is equal to the exact value of the oscillator's zero-point energy.

Problem 4.9
Find the energy levels of a particle of mass m moving in a one-dimensional potential:

$$V(x) = \begin{cases} +\infty, & x \le 0, \\ \frac{1}{2}m\omega^2 x^2, & x > 0. \end{cases}$$

Solution
This is an asymmetric harmonic oscillator potential in which the particle moves only in the region $x > 0$. The only acceptable solutions are those for which the wave function vanishes at $x = 0$. These solutions must be those of an ordinary (symmetric) harmonic oscillator that have odd parity, since the wave functions corresponding to the symmetric harmonic oscillator are either even (n even) or odd (n odd), and only the odd solutions vanish at the origin, $\psi_{2n+1}(0) = 0$ ($n = 0, 1, 2, 3, \ldots$). Therefore, the energy levels of this asymmetric potential must be given by those corresponding to the odd n energy levels of the symmetric potential, i.e.,

$$E_n = \left[(2n+1) + \frac{1}{2}\right]\hbar\omega = \left(2n + \frac{3}{2}\right)\hbar\omega \qquad (n = 0, 1, 2, 3, \ldots). \tag{4.297}$$

Problem 4.10
Consider the box potential

$$V(x) = \begin{cases} 0, & 0 < x < a, \\ \infty, & \text{elsewhere.} \end{cases}$$

(a) Estimate the energies of the ground state as well as those of the first and the second excited states for (i) an electron enclosed in a box of size $a = 10^{-10}$ m (express your answer in electron volts; you may use these values: $\hbar c = 200\,\text{MeV fm}$, $m_e c^2 = 0.5\,\text{MeV}$); (ii) a 1 g metallic sphere which is moving in a box of size $a = 10\,\text{cm}$ (express your answer in joules).

(b) Discuss the importance of the quantum effects for both of these two systems.

(c) Use the uncertainty principle to estimate the velocities of the electron and the metallic sphere.

Solution
The energy of a particle of mass m in a box having perfectly rigid walls is given by

$$E_n = \frac{n^2\hbar^2}{8ma^2}, \qquad n = 1, 2, 3, \ldots, \tag{4.298}$$

where a is the size of the box.

(a) (i) For the electron in the box of size 10^{-10} m, we have

$$E_n = \frac{\hbar^2 c^2}{m_e c^2 a^2}\frac{4\pi^2 n^2}{8} \equiv \frac{4\times 10^4\,(\text{MeV fm})^2}{0.5\,\text{MeV}\times 10^{10}\,\text{fm}^2}\frac{\pi^2}{2}n^2 = 4\pi^2 n^2\,\text{eV} \simeq 39n^2\,\text{eV}. \tag{4.299}$$

Hence $E_1 = 39\,\text{eV}$, $E_2 = 156\,\text{eV}$, and $E_3 = 351\,\text{eV}$.

(ii) For the sphere in the box of side 10 cm we have

$$E_n = \frac{(6.6\times 10^{-34}\,\text{J s})^2}{8\times 10^{-3}\,\text{kg}\times 10^{-2}\,\text{m}^2}n^2 = 5.45\times 10^{-63}n^2\,\text{J} \tag{4.300}$$

Hence $E_1 = 5.45\times 10^{-63}$ J, $E_2 = 21.8\times 10^{-63}$ J, and $E_3 = 49.1\times 10^{-63}$ J.

(b) The differences between the energy levels are

$$(E_2 - E_1)_{electron} = 117\,\text{eV}, \qquad (E_3 - E_2)_{electron} = 195\,\text{eV}, \tag{4.301}$$

$$(E_2 - E_1)_{sphere} = 16.35\times 10^{-63}\,\text{J}, \qquad (E_3 - E_2)_{sphere} = 27.3\times 10^{-63}\,\text{J}. \tag{4.302}$$

These results show that:

- The spacings between the energy levels of the electron are quite large; the levels are far apart from each other. Thus, the quantum effects are important.
- The energy levels of the sphere are practically indistinguishable; the spacings between the levels are negligible. The energy spectrum therefore forms a continuum; hence the quantum effects are not noticeable for the sphere.

(c) According to the uncertainty principle, the speed is proportional to $v \sim \hbar/(ma)$. For the electron, the typical distances are atomic, $a \simeq 10^{-10}$ m; hence

$$v \sim \frac{\hbar c}{mc^2 a}c \sim \frac{200\,\text{MeV fm}}{0.5\,\text{MeV}\times 10^5\,\text{fm}}c \simeq 4\times 10^{-3}c = 1.2\times 10^6\,\text{m s}^{-1}, \tag{4.303}$$

where c is the speed of light. The electron therefore moves quite fast; this is expected since we have confined the electron to move within a small region.

For the sphere, the typical distances are in the range of 1 cm:

$$v \sim \frac{\hbar}{ma} \sim \frac{6.6\times 10^{-34}\,\text{J s}}{10^{-3}\,\text{kg}\times 10^{-2}\,\text{m}} \simeq 6.6\times 10^{-29}\,\text{m s}^{-1} \tag{4.304}$$

At this speed the sphere is practically at rest.

Problem 4.11

(a) Verify that the matrices representing the operators $\hat{X}$ and $\hat{P}$ in the N-space for a harmonic oscillator obey the correct commutation relation $[\hat{X}, \hat{P}] = i\hbar$.

(b) Show that the energy levels of the harmonic oscillator can be obtained by inserting the matrices of $\hat{X}$ and $\hat{P}$ into the Hamiltonian $\hat{H} = \hat{P}^2/(2m) + \frac{1}{2}m\omega^2\hat{X}^2$.

Solution

(a) Using the matrices of $\hat{X}$ and $\hat{P}$ in (4.194) and (4.195), we obtain

$$\hat{X}\hat{P} = i\frac{\hbar}{2}\begin{pmatrix} 1 & 0 & -\sqrt{2} & \cdots \\ 0 & 1 & 0 & \cdots \\ \sqrt{2} & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad \hat{P}\hat{X} = i\frac{\hbar}{2}\begin{pmatrix} -1 & 0 & -\sqrt{2} & \cdots \\ 0 & -1 & 0 & \cdots \\ \sqrt{2} & 0 & -1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}; \tag{4.305}$$

hence

$$\hat{X}\hat{P} - \hat{P}\hat{X} = i\hbar\begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{4.306}$$

or $[\hat{X}, \hat{P}] = i\hbar I$, where I is the unit matrix.

(b) Again, using the matrices of $\hat{X}$ and $\hat{P}$ in (4.194) and (4.195), we can verify that

$$\hat{X}^2 = \frac{\hbar}{2m\omega}\begin{pmatrix} 1 & 0 & \sqrt{2} & \cdots \\ 0 & 3 & 0 & \cdots \\ \sqrt{2} & 0 & 5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad \hat{P}^2 = -\frac{m\hbar\omega}{2}\begin{pmatrix} -1 & 0 & \sqrt{2} & \cdots \\ 0 & -3 & 0 & \cdots \\ \sqrt{2} & 0 & -5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}; \tag{4.307}$$

hence

$$\begin{aligned} \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2\hat{X}^2 &= \frac{1}{2m}\left(-\frac{m\hbar\omega}{2}\right)\begin{pmatrix} -1 & 0 & \sqrt{2} & \cdots \\ 0 & -3 & 0 & \cdots \\ \sqrt{2} & 0 & -5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} + \frac{1}{2}m\omega^2\left(\frac{\hbar}{2m\omega}\right)\begin{pmatrix} 1 & 0 & \sqrt{2} & \cdots \\ 0 & 3 & 0 & \cdots \\ \sqrt{2} & 0 & 5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \\ &= \frac{\hbar\omega}{4}\begin{pmatrix} 1 & 0 & -\sqrt{2} & \cdots \\ 0 & 3 & 0 & \cdots \\ -\sqrt{2} & 0 & 5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} + \frac{\hbar\omega}{4}\begin{pmatrix} 1 & 0 & \sqrt{2} & \cdots \\ 0 & 3 & 0 & \cdots \\ \sqrt{2} & 0 & 5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \\ &= \frac{\hbar\omega}{2}\begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 3 & 0 & \cdots \\ 0 & 0 & 5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \end{aligned} \tag{4.308}$$

The form of this matrix is similar to the result we obtain from an analytical treatment, $E_n = \hbar\omega(2n + 1)/2$, since

$$H_{n'n} = \langle n'|\hat{H} \mid n\rangle = \frac{\hbar\omega}{2}(2n + 1)\delta_{n'n}. \tag{4.309}$$

Problem 4.12
Calculate the probability of finding a particle in the classically forbidden region of a harmonic oscillator for the states $n = 0, 1, 2, 3, 4$. Are these results compatible with their classical counterparts?

Solution
The classical turning points are defined by $E_n = V(x_n)$ or by $\hbar\omega(n + \frac{1}{2}) = \frac{1}{2}m\omega^2 x_n^2$; that is, $x_n = \pm\sqrt{\hbar/(m\omega)}\sqrt{2n+1}$. Thus, the probability of finding a particle in the classically forbidden region for a state $\psi_n(x)$ is

$$P_n = \int_{-\infty}^{-|x_n|} |\psi_n(x)|^2 dx + \int_{|x_n|}^{+\infty} |\psi_n(x)|^2 dx = 2\int_{|x_n|}^{+\infty} |\psi_n(x)|^2 dx, \tag{4.310}$$

where $\psi_n(x)$ is given in (4.186), $\psi_n(x) = 1/\sqrt{\sqrt{\pi}2^n n! x_0} e^{-x^2/2x_0^2} H_n(x/x_0)$, where x_0 is given by $x_0 = \sqrt{\hbar/(m\omega)}$. Using the change of variable $y = x/x_0$, we can rewrite P_n as

$$P_n = \frac{2}{\sqrt{\pi}2^n n!}\int_{\sqrt{2n+1}}^{+\infty} e^{-y^2} H_n^2(y)\, dy, \tag{4.311}$$

where the Hermite polynomials $H_n(y)$ are listed in (4.129). The integral in (4.311) can be evaluated only numerically. Using the numerical values

$$\int_1^\infty e^{-y^2} dy = 0.1394, \qquad \int_{\sqrt{3}}^\infty y^2 e^{-y^2} dy = 0.0495, \tag{4.312}$$

$$\int_{\sqrt{5}}^\infty \left(4y^2 - 2\right)^2 e^{-y^2} dy = 0.6740, \qquad \int_{\sqrt{7}}^\infty \left(8y^3 - 12y\right)^2 e^{-y^2} dy = 3.6363, \tag{4.313}$$

$$\int_{\sqrt{9}}^\infty \left(16y^4 - 48y^2 + 12\right)^2 e^{-y^2} dx = 26.86, \tag{4.314}$$

we obtain

$$P_0 = 0.1573, \qquad P_1 = 0.1116, \qquad P_2 = 0.095\,069, \tag{4.315}$$

$$P_3 = 0.085\,48, \qquad P_4 = 0.078\,93. \tag{4.316}$$

This shows that the probability decreases as n increases, so it would be very small for very large values of n. It is therefore unlikely to find the particle in the classically forbidden region when the particle is in a very highly excited state. This is what we expect, since the classical approximation is recovered in the limit of high values of n.

Problem 4.13
Consider a particle of mass m moving in the following potential

$$V(x) = \begin{cases} \infty, & x \le 0, \\ -V_0, & 0 < x < a, \\ 0, & x \ge a, \end{cases}$$

where $V_0 > 0$.

(a) Find the wave function.

(b) Show how to obtain the energy eigenvalues from a graph.

(c) Calculate the minimum value of V_0 (in terms of m, a, and $\hbar$) so that the particle will have one bound state; then calculate it for two bound states. From these two results, try to obtain the lowest value of V_0 so that the system has n bound states.

Solution

(a) As shown in Figure 4.16, the wave function in the region $x < 0$ is zero, $\psi(x) = 0$. In the region $x > 0$ the Schrödinger equation for the bound state solutions, $-V_0 < E < 0$, is given by

$$\left(\frac{d^2}{dx^2} + k_1^2\right)\psi_1(x) = 0 \qquad (0 < x < a), \tag{4.317}$$

$$\left(\frac{d^2}{dx^2} - k_2^2\right)\psi_2(x) = 0 \qquad (x > a), \tag{4.318}$$

where $k_1^2 = 2m(V_0+E)/\hbar^2$ and $k_2^2 = -2mE/\hbar^2$. On one hand, the solution of (4.317) is oscillatory, $\psi_1(x) = A \sin k_1x + B \cos k_1x$, but since $\psi_1(0) = 0$ we must have $B = 0$. On the other hand, eliminating the physically unacceptable solutions which grow exponentially for large values of x, the solution of (4.318) is $\psi_2(x) = Ce^{-k_2x}$. Thus, the wave function is given by

$$\psi(x) = \begin{cases} 0, & x < 0, \\ \psi_1(x) = A \sin k_1x, & 0 < x < 0, \\ \psi_2(x) = Ce^{-k_2x}, & x > a. \end{cases} \tag{4.319}$$

(b) To determine the eigenvalues, we need to use the boundary conditions at $x = a$. The condition $\psi_1(a) = \psi_2(a)$ yields

$$A \sin k_1a = Ce^{-k_2a}, \tag{4.320}$$

while the continuity of the first derivative, $\psi_1'(a) = \psi_2'(a)$, leads to

$$Ak_1 \cos k_1a = -Ck_2e^{-k_2a}. \tag{4.321}$$

Dividing (4.321) by (4.320) we obtain

$$k_1a \cot k_1a = -k_2a. \tag{4.322}$$

Since $k_1^2 = 2m(V_0 + E)/\hbar^2$ and $k_2^2 = -2mE/\hbar^2$, we have

$$(k_1a)^2 + (k_2a)^2 = \gamma^2, \tag{4.323}$$

where $\gamma = \sqrt{2mV_0}a/\hbar$.

The transcendental equations (4.322) and (4.323) can be solved graphically. As shown in Figure 4.16, the energy levels are given by the intersection of the circular curve $(k_1a)^2 + (k_2a)^2 = \gamma^2$ with $k_1a \cot k_1a = -k_2a$.

(c) If $\pi/2 < \gamma < 3\pi/2$ there will be only one bound state, the ground state $n = 1$, for there is only one crossing between the curves $(k_1a)^2 + (k_2a)^2 = \gamma^2$ and $k_1a \cot k_1a = -k_2a$. The lowest value of V_0 that yields a single bound state is given by the relation $\gamma = \pi/2$, which leads to $2ma^2V_0/\hbar^2 = \pi^2/4$ or to

$$V_0 = \frac{\pi^2\hbar^2}{8ma^2}. \tag{4.324}$$

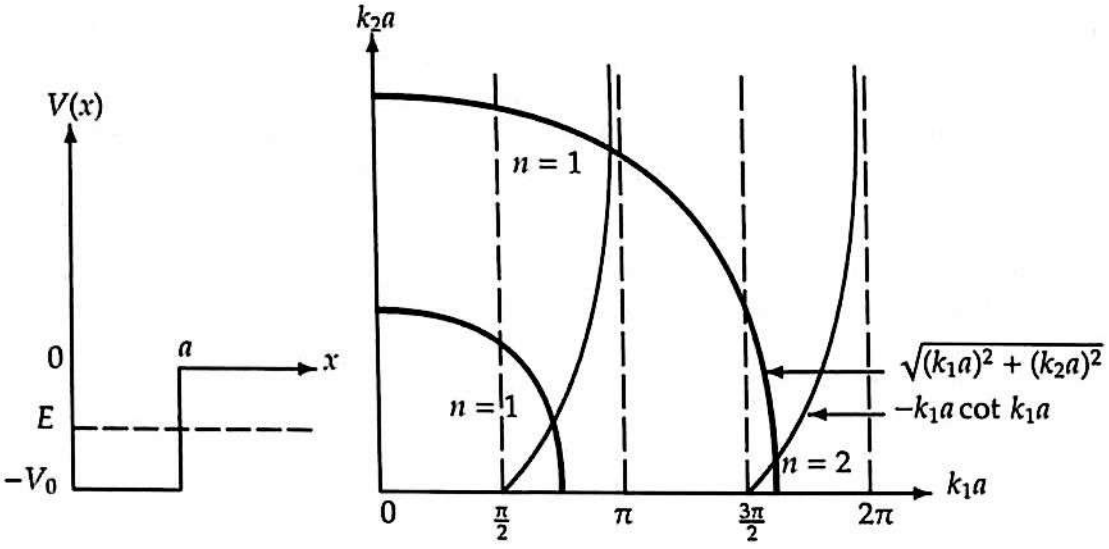


Figure 4.16: Potential $V(x)$ (left curve); the energy levels of $V(x)$ are given graphically by the intersection of the circular curve $\sqrt{(k_1a)^2 + (k_2a)^2}$ with $-k_1a \cot k_1a$ (right curve).

Similarly, if $3\pi/2 < \gamma < 5\pi/2$ there will be two crossings between $(k_1a)^2 + (k_2a)^2 = \gamma^2$ and $k_1a \cot k_1a = -k_2a$. Thus, there will be two bound states: the ground state, $n = 1$, and the first excited state, $n = 2$. The lowest value of V_0 that yields two bound states corresponds to $2ma^2V_0/\hbar^2 = 9\pi^2/4$ or to

$$V_0 = \frac{9\pi^2\hbar^2}{8ma^2}. \tag{4.325}$$

We may thus infer the following general result. If $n\pi - \pi/2 < \gamma < n\pi + \pi/2$, there will be n crossings and hence n bound states:

$$n\pi - \frac{\pi}{2} < \frac{\sqrt{2mV_0}}{\hbar}a < n\pi + \frac{\pi}{2} \quad \Longrightarrow \quad \text{there are } n \text{ bound states.} \tag{4.326}$$

The lowest value of V_0 giving n bound states is

$$V_0 = \frac{\pi^2\hbar^2}{8ma^2}(2n-1)^2. \tag{4.327}$$

Problem 4.14

(a) Assuming the potential seen by a neutron in a nucleus to be schematically represented by a one-dimensional, infinite rigid walls potential of length 10 fm, estimate the minimum kinetic energy of the neutron.

(b) Estimate the minimum kinetic energy of an electron bound within the nucleus described in (a). Can an electron be confined in a nucleus? Explain.

Solution

The energy of a particle of mass m in a one-dimensional box potential having perfectly rigid walls is given by

$$E_n = \frac{\pi^2\hbar^2}{2ma^2}n^2, \qquad n = 1, 2, 3, \ldots, \tag{4.328}$$

where a is the size of the box.

(a) Assuming the neutron to be nonrelativistic (i.e., its energy $E \ll m_n c^2$), the lowest energy the neutron can have in a box of size $a = 10\,\text{fm}$ is

$$E_{min} = \frac{\pi^2\hbar^2}{2m_n a^2} = \frac{\pi^2(\hbar^2c^2)}{2(m_nc^2)a^2} \simeq 2.04\,\text{MeV}, \tag{4.329}$$

where we have used the fact that the rest mass energy of a neutron is $m_n c^2 \simeq 939.57\,\text{MeV}$ and $\hbar c \simeq 197.3\,\text{MeV fm}$. Indeed, we see that $E_{min} \ll m_n c^2$.

(b) The minimum energy of a (nonrelativistic) electron moving in a box of size $a = 10\,\text{fm}$ is given by

$$E_{min} = \frac{\pi^2\hbar^2}{2m_e a^2} = \frac{\pi^2(\hbar^2c^2)}{2(m_ec^2)a^2} \simeq 3755.45\,\text{MeV}. \tag{4.330}$$

The rest mass energy of an electron is $m_e c^2 \simeq 0.511\,\text{MeV}$, so this electron is ultra-relativistic since $E_{min} \gg m_e c^2$. It implies that an electron with this energy cannot be confined within such a nucleus.

Problem 4.15

(a) Calculate the expectation value of the operator $\hat{X}^4$ in the N-representation with respect to the state $|\,n\rangle$ (i.e., $\langle n \mid \hat{X}^4 \mid n\rangle$).

(b) Use the result of (a) to calculate the energy E_n for a particle whose Hamiltonian is $\hat{H} = \hat{P}^2/(2m) + \frac{1}{2}m\omega^2\hat{X}^2 - \lambda\hat{X}^4$.

Solution

(a) Since $\sum_{m=0}^{\infty} |m\rangle\langle m| = 1$ we can write the expectation value of $\hat{X}^4$ as

$$\langle n \mid \hat{X}^4 \mid n\rangle = \sum_{m=0}^{\infty} \langle n \mid \hat{X}^2 | m\rangle\langle m|\hat{X}^2 \mid n\rangle = \sum_{m=0}^{\infty} \left|\langle m|\hat{X}^2 \mid n\rangle\right|^2. \tag{4.331}$$

Now since

$$\hat{X}^2 = \frac{\hbar}{2m\omega}\left(\hat{a}^2 + \hat{a}^{\dagger 2} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right) = \frac{\hbar}{2m\omega}\left(\hat{a}^2 + \hat{a}^{\dagger 2} + 2\hat{a}^\dagger\hat{a} + 1\right), \tag{4.332}$$

the only terms $\langle m|\hat{X}^2 \mid n\rangle$ that survive are

$$\langle n \mid \hat{X}^2 \mid n\rangle = \frac{\hbar}{2m\omega}\langle n \mid 2\hat{a}^\dagger\hat{a} + 1 \mid n\rangle = \frac{\hbar}{2m\omega}(2n+1), \tag{4.333}$$

$$\langle n-2 \mid \hat{X}^2 \mid n\rangle = \frac{\hbar}{2m\omega}\langle n-2 \mid \hat{a}^2 \mid n\rangle = \frac{\hbar}{2m\omega}\sqrt{n(n-1)}, \tag{4.334}$$

$$\langle n+2 \mid \hat{X}^2 \mid n\rangle = \frac{\hbar}{2m\omega}\langle n+2 \mid \hat{a}^{\dagger 2} \mid n\rangle = \frac{\hbar}{2m\omega}\sqrt{(n+1)(n+2)}. \tag{4.335}$$

Thus

$$\begin{aligned}\langle n \mid \hat{X}^4 \mid n\rangle &= \left|\langle n \mid \hat{X}^2 \mid n\rangle\right|^2 + \left|\langle n-2 \mid \hat{X}^2 \mid n\rangle\right|^2 + \left|\langle n+2 \mid \hat{X}^2 \mid n\rangle\right|^2 \\ &= \frac{\hbar^2}{4m^2\omega^2}\left[(2n+1)^2 + n(n-1) + (n+1)(n+2)\right] \\ &= \frac{\hbar^2}{4m^2\omega^2}\left(6n^2 + 6n + 3\right).\end{aligned} \tag{4.336}$$

(b) Using (4.336), and since the Hamiltonian can be expressed in terms of the harmonic oscillator, $\hat{H} = \hat{H}_{HO} - \lambda\hat{X}^4$, we immediately obtain the particle energy:

$$E_n = \langle n \mid \hat{H}_{HO} \mid n\rangle - \lambda\langle n \mid \hat{X}^4 \mid n\rangle = \hbar\omega\left(n + \frac{1}{2}\right) - \frac{\lambda\hbar^2}{4m^2\omega^2}\left(6n^2 + 6n + 3\right). \tag{4.337}$$

Problem 4.16
Find the energy levels and the wave functions of two harmonic oscillators of masses m_1 and m_2, having identical frequencies ω, and coupled by the interaction $\frac{1}{2}k(\hat{X}_1 - \hat{X}_2)^2$.

Solution
This problem reduces to finding the eigenvalues for the Hamiltonian

$$\begin{aligned}\hat{H} &= \hat{H}_1 + \hat{H}_2 + \frac{1}{2}K(\hat{X}_1 - \hat{X}_2)^2 \\ &= \frac{1}{2m_1}\hat{p}_1^2 + \frac{1}{2}m_1\omega^2\hat{X}_1^2 + \frac{1}{2m_2}\hat{p}_2^2 + \frac{1}{2}m_2\omega^2\hat{X}_2^2 + \frac{1}{2}K(\hat{X}_1 - \hat{X}_2)^2.\end{aligned} \tag{4.338}$$

This is a two-particle problem. As in classical mechanics, it is more convenient to describe the dynamics of a two-particle system in terms of the center of mass (CM) and relative motions. For this, let us introduce the following operators:

$$\hat{P} = \hat{p}_1 + \hat{p}_2, \qquad \hat{X} = \frac{m_1\hat{x}_1 + m_2\hat{x}_2}{M}, \tag{4.339}$$

$$\hat{p} = \frac{m_2\hat{p}_1 - m_1\hat{p}_2}{M}, \qquad \hat{x} = \hat{x}_1 - \hat{x}_2, \tag{4.340}$$

where $M = m_1 + m_2$ and $\mu = m_1m_2/(m_1 + m_2)$ is the reduced mass; $\hat{P}$ and $\hat{X}$ pertain to the CM; $\hat{p}$ and $\hat{x}$ pertain to the relative motion. These relations lead to

$$\hat{p}_1 = \frac{m_1}{M}\hat{P} + \hat{p}, \qquad \hat{p}_2 = \frac{m_2}{M}\hat{P} - \hat{p}, \tag{4.341}$$

$$\hat{x}_1 = \frac{m_2}{M}\hat{x} + \hat{X}, \qquad \hat{x}_2 = -\frac{m_1}{M}\hat{x} + \hat{X}. \tag{4.342}$$

Note that the sets (X, P) and (x, p) are conjugate variables separately: $[\hat{X}, \hat{P}] = i\hbar$, $[\hat{x}, \hat{p}] = i\hbar$, $[\hat{X}, \hat{p}] = [\hat{x}, \hat{P}] = 0$. Taking $\hat{p}_1$, $\hat{p}_2$, $\hat{x}_1$, and $\hat{x}_2$ of (4.341) and (4.342) and inserting them into (4.338), we obtain

$$\begin{aligned}\hat{H} &= \frac{1}{2m_1}\left(\frac{m_1}{M}\hat{P} + \hat{p}\right)^2 + \frac{1}{2}m_1\omega^2\left(\frac{m_2}{M}\hat{x} + \hat{X}\right)^2 \\ &\quad + \frac{1}{2m_2}\left(\frac{m_2}{M}\hat{P} - \hat{p}\right)^2 + \frac{1}{2}m_2\omega^2\left(-\frac{m_1}{M}\hat{x} + \hat{X}\right)^2 + \frac{1}{2}K\hat{x}^2 \\ &= \hat{H}_{CM} + \hat{H}_{rel},\end{aligned} \tag{4.343}$$

where

$$\hat{H}_{CM} = \frac{1}{2M}\hat{P}^2 + \frac{1}{2}M\omega^2\hat{X}^2, \qquad \hat{H}_{rel} = \frac{1}{2\mu}\hat{p}^2 + \frac{1}{2}\mu\Omega^2\hat{x}^2, \tag{4.344}$$

with $\Omega^2 = \omega^2 + k/\mu$. We have thus reduced the Hamiltonian of these two coupled harmonic oscillators to the sum of two independent harmonic oscillators, one with frequency ω and mass M and the other of mass μ and frequency $\Omega = \sqrt{\omega^2 + k/\mu}$. That is, by introducing the CM and relative motion variables, we have managed to eliminate the coupled term from the Hamiltonian.

The energy levels of this two-oscillator system can be inferred at once from the suggestive Hamiltonians of (4.344):

$$E_{n_1 n_2} = \hbar\omega\left(n_1 + \frac{1}{2}\right) + \hbar\Omega\left(n_2 + \frac{1}{2}\right). \tag{4.345}$$

The states of this two-particle system are given by the product of the two states $|N\rangle = |n_1\rangle|n_2\rangle$; hence the total wave function, $\psi_n(X, x)$, is equal to the product of the center of mass wave function, $\psi_{n_1}(X)$, and the wave function of the relative motion, $\psi_{n_2}(x)$: $\psi_n(X, x) = \psi_{n_1}(X)\psi_{n_2}(x)$. Note that both of these wave functions are harmonic oscillator functions whose forms can be found in (4.186):

$$\psi_n(X, x) = \frac{1}{\sqrt{\pi}\sqrt{2^{n_1}2^{n_2}n_1!n_2!x_{0_1}x_{0_2}}} e^{-X^2/2x_{0_1}^2} e^{-x^2/2x_{0_2}^2} H_{n_1}\left(\frac{X}{x_{0_1}}\right) H_{n_2}\left(\frac{x}{x_{0_2}}\right), \tag{4.346}$$

where $n = (n_1, n_2)$, $x_{0_1} = \sqrt{\hbar/(M\omega)}$, and $x_{0_2} = \sqrt{\hbar/(\mu\Omega)}$.

Problem 4.17

Consider a particle of mass m and charge q moving under the influence of a one-dimensional harmonic oscillator potential. Assume it is placed in a constant electric field $\mathcal{E}$. The Hamiltonian of this particle is therefore given by $\hat{H} = \hat{P}^2/(2m) + \frac{1}{2}m\omega^2\hat{X}^2 - q\mathcal{E}\hat{X}$. Derive the energy expression and the wave function of the nth excited state.

Solution

To find the eigenenergies of the Hamiltonian

$$\hat{H} = \frac{1}{2m}\hat{P}^2 + \frac{1}{2}m\omega^2\hat{X}^2 - q\mathcal{E}\hat{X}, \tag{4.347}$$

it is convenient to use the change of variable $y = \hat{X} - q\mathcal{E}/(m\omega^2)$. Thus the Hamiltonian becomes

$$\hat{H} = \frac{1}{2m}\hat{P}^2 + \frac{1}{2}m\omega^2\hat{y}^2 - \frac{q^2\mathcal{E}^2}{2m\omega^2}. \tag{4.348}$$

Since the term $q^2\mathcal{E}^2/(2m\omega^2)$ is a mere constant and $\hat{P}^2/(2m) + \frac{1}{2}m\omega^2\hat{y}^2 = \hat{H}_{HO}$ has the structure of a harmonic oscillator Hamiltonian, we can easily infer the energy levels:

$$E_n = \langle n \mid \hat{H} \mid n\rangle = \hbar\omega\left(n + \frac{1}{2}\right) - \frac{q^2\mathcal{E}^2}{2m\omega^2}. \tag{4.349}$$

The wave function is given by $\psi_n(y) = \psi_n(x - q\mathcal{E}/(m\omega^2))$, where $\psi_n(y)$ is given in (4.186):

$$\psi_n(y) = \frac{1}{\sqrt{\sqrt{\pi}2^n n! x_0}} e^{-y^2/2x_0^2} H_n\left(\frac{y}{x_0}\right). \tag{4.350}$$

Problem 4.18
Consider a particle of mass m that is bouncing vertically and elastically on a smooth reflecting floor in the Earth's gravitational field

$$V(z) = \begin{cases} mgz, & z > 0, \\ +\infty, & z \le 0, \end{cases}$$

where g is a constant (the acceleration due to gravity). Find the energy levels and wave function of this particle.

Solution
We need to solve the Schrödinger equation with the boundary conditions $\psi(0) = 0$ and $\psi(+\infty) = 0$:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(z)}{dz^2} + mgz\psi(z) = E\psi(z) \implies \frac{d^2\psi(z)}{dz^2} - \frac{2m}{\hbar^2}(mgz - E)\psi(z) = 0. \tag{4.351}$$

With the change of variable $x = (\hbar^2/(2m^2g))^{2/3}(2m/\hbar^2)(mgz - E)$, we can reduce this equation to

$$\frac{d^2\phi(x)}{dx^2} - x\phi(x) = 0. \tag{4.352}$$

This is a standard differential equation; its solution (which vanishes at $x \to +\infty$, i.e., $\phi(+\infty) = 0$) is given by

$$\phi(x) = B\mathrm{Ai}(x) \qquad \text{where} \qquad \mathrm{Ai}(x) = \frac{1}{\pi}\int_0^\infty \cos\left(\frac{1}{3}t^3 + xt\right)dt, \tag{4.353}$$

where $\mathrm{Ai}(x)$ is called the *Airy function*.

When $z = 0$ we have $x = -(2/(mg^2\hbar^2))^{1/3}E$. The boundary condition $\psi(0) = 0$ yields $\phi[-(2/(mg^2\hbar^2))^{1/3}E] = 0$ or $\mathrm{Ai}[-(2/(mg^2\hbar^2))^{1/3}E] = 0$. The Airy function has zeros only at certain values of R_n: $\mathrm{Ai}(R_n) = 0$ with $n = 0, 1, 2, 3, \ldots$. The roots R_n of the Airy function can be found in standard tables. For instance, the first few roots are $R_0 = -2.338$, $R_1 = -4.088$, $R_2 = -5.521$, $R_3 = -6.787$.

The boundary condition $\psi(0) = 0$ therefore gives a *discrete* set of energy levels which can be expressed in terms of the roots of the Airy function:

$$\mathrm{Ai}\left[-\left(\frac{2}{mg^2\hbar^2}\right)^{1/3}E\right] = 0 \implies -\left(\frac{2}{mg^2\hbar^2}\right)^{1/3}E_n = R_n; \tag{4.354}$$

hence

$$E_n = -\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3}R_n, \qquad \psi_n(z) = B_n\mathrm{Ai}\left[-\left(\frac{2m^2g^2}{\hbar^2}\right)^{1/3}z - R_n\right]. \tag{4.355}$$

The first few energy levels are

$$E_0 = 2.338\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3}, \qquad E_1 = 4.088\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3}, \tag{4.356}$$

$$E_2 = 5.521\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3}. \qquad E_3 = 6.787\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3}. \tag{4.357}$$

4.11 Exercises

Exercise 4.1
A particle of mass m is subjected to a potential

$$V(x) = \begin{cases} 0, & |x| < a/2, \\ \infty, & |x| > a/2. \end{cases}$$

(a) Find the ground, first, and second excited state wave functions.
(b) Find expressions for E_1, E_2, and E_3.
(c) Plot the probability densities $P_2(x,t)$ and $P_3(x,t)$.
(d) Find $\langle X\rangle_2$, $\langle X\rangle_3$, $\langle P\rangle_2$, and $\langle P\rangle_3$.
(e) Evaluate $\Delta x \Delta p$ for the states $\psi_2(x,t)$ and $\psi_3(x,t)$.

Exercise 4.2
Consider a system whose wave function at $t = 0$ is

$$\psi(x,0) = \frac{3}{\sqrt{30}}\phi_0(x) + \frac{4}{\sqrt{30}}\phi_1(x) + \frac{1}{\sqrt{6}}\phi_4(x),$$

where $\phi_n(x)$ is the wave function of the nth excited state of an infinite square well potential of width a and whose energy is $E_n = \pi^2\hbar^2 n^2/(2ma^2)$.

(a) Find the average energy of this system.

(b) Find the state $\psi(x,t)$ at a later time t and the average value of the energy. Compare the result with the value obtained in (a).

Exercise 4.3
An electron with a kinetic energy of 10 eV at large negative values of x is moving from left to right along the x-axis. The potential energy is

$$V(x) = \begin{cases} 0 & (x \le 0), \\ 20\,\text{eV} & (x > 0). \end{cases}$$

(a) Write the time-independent Schrödinger equation in the regions $x \le 0$ and $x > 0$.
(b) Describe the shapes for $\psi(x)$ for $x \le 0$ and $x > 0$.
(c) Calculate the electron wavelength (in meters) in $-20\,\text{m} < x < -10\,\text{m}$ and $x > 10\,\text{m}$.
(d) Write down the boundary conditions at $x = 0$.

(e) Calculate the ratio of the probabilities for finding the electron near $x = 10^{-10}\,\text{m}$ and $x = 0$.

Exercise 4.4
A particle is moving in the potential well

$$V(x) = \begin{cases} 0, & -a \le x \le -b, \\ V_0, & -b \le x \le b, \\ 0, & b \le x \le a, \\ +\infty & \text{elsewhere} \end{cases},$$

where V_0 is positive. In this problem consider $E < V_0$. Let $\psi_1(x)$ and $\psi_2(x)$ represent the two lowest energy solutions of the Schrödinger equation; call their energies E_1 and E_2, respectively.

(a) Calculate E_1 and E_2 in units of eV for the case where $mc^2 = 1\,\text{GeV}$, $a = 10^{-14}\,\text{m}$, and $b = 0.4 \times 10^{-14}\,\text{m}$; take $\hbar c \simeq 200\,\text{MeV fm}$.

(b) A particular solution of the Schrödinger equation can be constructed by superposing $\psi_1(x)e^{iE_1t/\hbar}$ and $\psi_2(x)e^{iE_2t/\hbar}$. Construct a wave packet ψ which at $t = 0$ is (almost) entirely to the left-hand side of the well and describe its motion in time; find the period of oscillations between the two terms of ψ.

Exercise 4.5
A particle moves in the potential

$$V(x) = \frac{\hbar^2}{2m}\left[\frac{4}{225}\sinh^2 x - \frac{2}{5}\cosh x\right].$$

(a) Sketch $V(x)$ and locate the position of the two minima.

(b) Show that $\psi(x) = (1 + 4\cosh x)\exp\left(-\frac{2}{15}\cosh x\right)$ is a solution of the time-independent Schrödinger equation for the particle. Find the corresponding energy level and indicate its position on the sketch of $V(x)$.

(c) Sketch $\psi(x)$ and show that it has the proper behavior at the classical turning points and in the classically forbidden regions.

Exercise 4.6
Show that for a particle of mass m which moves in a one-dimensional infinite potential well of length a, the uncertainties product $\Delta x_n \Delta p_n$ is given by $\Delta x_n \Delta p_n \simeq n\pi\hbar/\sqrt{12}$.

Exercise 4.7
A particle of mass m is moving in an infinite potential well

$$V(x) = \begin{cases} V_0, & 0 < x < a, \\ \infty, & \text{elsewhere.} \end{cases}$$

(a) Solve the Schrödinger equation and find the energy levels and the corresponding normalized wave functions.

(b) Calculate $\langle\hat{X}\rangle_5$, $\langle\hat{P}\rangle_5$, $\langle\hat{X}^2\rangle_5$, and $\langle\hat{P}^2\rangle_5$ for the fourth excited state and infer the value of $\Delta x \Delta p$.

Exercise 4.8
Consider the potential step

$$V(x) = \begin{cases} 6\,\text{eV}, & x < 0, \\ 0, & x > 0. \end{cases}$$

(a) An electron of energy 8 eV is moving from left to right in this potential. Calculate the probability that the electron will (i) continue moving along its initial direction after reaching the step and (ii) get reflected at the potential step.

(b) Now suppose the electron is moving from right to left with an energy 3 eV. (i) Estimate the order of magnitude of the distance the electron can penetrate the barrier. (ii) Repeat part (i) for a 70 kg person initially moving at $4\,\text{m s}^{-1}$ and running into a wall which can be represented by a potential step of height equal to four times this person's energy before reaching the step.

Exercise 4.9
Consider a system whose wave function at time $t = 0$ is given by

$$\psi(x,0) = \frac{5}{\sqrt{50}}\phi_0(x) + \frac{4}{\sqrt{50}}\phi_1(x) + \frac{3}{\sqrt{50}}\phi_2(x),$$

where $\phi_n(x)$ is the wave function of the nth excited state for a harmonic oscillator of energy $E_n = \hbar\omega(n + 1/2)$.

(a) Find the average energy of this system.

(b) Find the state $\psi(x,t)$ at a later time t and the average value of the energy; compare the result with the value obtained in (a).

(c)Find the expectation value of the operator $\hat{X}$ with respect to the state $\psi(x,t)$ (i.e., find $\langle\psi(x,t)|\hat{X}|\psi(x,t)\rangle$).

Exercise 4.10
Calculate $\langle n \mid \hat{X}^2 \mid m\rangle$ and $\langle m \mid \hat{X}^4 \mid n\rangle$ in the N-representation; $\mid n\rangle$ and $\mid m\rangle$ are harmonic oscillator states.

Exercise 4.11
Consider the dimensionless Hamiltonian $\hat{H} = \frac{1}{2}\hat{P}^2 + \frac{1}{2}\hat{X}^2$, with $\hat{P} = -id/dx$.

(a) Show that the wave functions $\psi_0(x) = e^{-x^2/2}/\sqrt{\sqrt{\pi}}$ and $\psi_1(x) = \sqrt{2/\sqrt{\pi}}xe^{-x^2/2}$ are eigenfunctions of $\hat{H}$ with eigenvalues 1/2 and 3/2, respectively.

(b) Find the values of the coefficients α and β such that

$$\psi_2(x) = \frac{1}{\sqrt{2\sqrt{\pi}}}\left(\alpha x^2 - 1\right)e^{-x^2/2} \quad \text{and} \quad \psi_3(x) = \frac{1}{\sqrt{6\sqrt{\pi}}}x\left(1 + \beta x^2\right)e^{-x^2/2}$$

are orthogonal to $\psi_0(x)$ and $\psi_1(x)$, respectively. Then show that $\psi_2(x)$ and $\psi_3(x)$ are eigenfunctions of $\hat{H}$ with eigenvalues 5/2 and 7/2, respectively.

Exercise 4.12
Consider the dimensionless Hamiltonian $\hat{H} = \frac{1}{2}\hat{P}^2 + \frac{1}{2}\hat{X}^2$ (with $\hat{P} = -id/dx$) whose wave function at time $t = 0$ is given by

$$\Psi(x,0) = \frac{1}{\sqrt{2}}\psi_0(x) + \frac{1}{\sqrt{8}}\psi_1(x) + \frac{1}{\sqrt{10}}\psi_2(x),$$

where $\psi_0(x) = \frac{1}{\sqrt{\sqrt{\pi}}}e^{-x^2/2}$, $\psi_1(x) = \sqrt{\frac{2}{\sqrt{\pi}}}xe^{-x^2/2}$, and $\psi_2(x) = \frac{1}{\sqrt{2\sqrt{\pi}}}\left(2x^2 - 1\right)e^{-x^2/2}$.

(a) Calculate $\Delta x_n \Delta p_n$ for $n = 0, 1$ where $\Delta x_n = \sqrt{\langle\psi_n|\hat{X}^2|\psi_n\rangle - \langle\psi_n|\hat{X}|\psi_n\rangle^2}$.

(b) Calculate $\hat{a}^\dagger\psi_0(x)$, $\hat{a}\psi_0(x)$, $\hat{a}^\dagger\psi_1(x)$, $\hat{a}\psi_1(x)$, and $\hat{a}\psi_2(x)$, where the operators $\hat{a}^\dagger$ and $\hat{a}$ are defined by $\hat{a} = (\hat{X} + d/dx)/\sqrt{2}$ and $\hat{a}^\dagger = (\hat{X} - d/dx)/\sqrt{2}$.

Exercise 4.13
Consider a particle of mass m that is moving in a one-dimensional infinite potential well with walls at $x = 0$ and $x = a$ which is initially (i.e., at $t = 0$) in the state

$$\psi(x,0) = \frac{1}{\sqrt{2}}\left[\phi_1(x) + \phi_3(x)\right],$$

where $\phi_1(x)$ and $\phi_3(x)$ are the ground and second excited states, respectively.

(a) What is the state vector $\psi(x,t)$ for $t > 0$ in the Schrödinger picture.

(b) Find the expectation values $\langle \hat{X} \rangle$, $\langle \hat{P} \rangle$, $\langle \hat{X}^2 \rangle$, and $\langle \hat{P}^2 \rangle$ with respect to $|\psi\rangle$.

(c) Evaluate $\Delta x \Delta p$ and verify that it satisfies the uncertainty principle.

Exercise 4.14

If the state of a particle moving in a one-dimensional harmonic oscillator is given by

$$|\psi\rangle = \frac{1}{\sqrt{17}}|0\rangle + \frac{3}{\sqrt{17}}|1\rangle - \frac{2}{\sqrt{17}}|2\rangle - \sqrt{\frac{3}{17}}|3\rangle,$$

where $|n\rangle$ represents the normalized nth energy eigenstate, find the expectation values of the number operator, $\hat{N}$, and of the Hamiltonian operator.

Exercise 4.15

Find the number of bound states and the corresponding energies for the finite square well potential when (a) $R = 7$ (i.e., $\sqrt{ma^2V_0/(2\hbar^2)} = 7$) and (b) $R = 3\pi$.

Exercise 4.16

A ball of mass $m = 0.2\,\text{kg}$ bouncing on a table located at $z = 0$ is subject to the potential

$$V(z) = \begin{cases} V_0 & (z < 0), \\ mgz & (z > 0), \end{cases}$$

where $V_0 = 3\,\text{J}$ and g is the acceleration due to gravity.

(a) Describe the spectrum of possible energies (i.e., continuous, discrete, or nonexistent) as E increases from large negative values to large positive values.

(b) Estimate the order of magnitude for the lowest energy state.

(c) Describe the general shapes of the wave functions $\psi_0(z)$ and $\psi_1(z)$ corresponding to the lowest two energy states and sketch the corresponding probability densities.

Exercise 4.17

Consider a particle of mass m moving in a one-dimensional harmonic oscillator potential, with $\hat{X} = \sqrt{\hbar/(2m\omega)}(\hat{a} + \hat{a}^\dagger)$ and $\hat{P} = i\sqrt{m\hbar\omega/2}(\hat{a}^\dagger - \hat{a})$.

(a) Calculate the product of the uncertainties in position and momentum for the particle in the fifth excited state, i.e., $(\Delta X \Delta P)_5$.

(b) Compare the result of (a) with the uncertainty product when the particle is in its lowest energy state. Explain why the two uncertainty products are different.

Exercise 4.18

A particle of mass m in an infinite potential well of length a has the following initial wave function at $t = 0$:

$$\psi(x,0) = \frac{2}{\sqrt{7a}}\sin\left(\frac{\pi x}{a}\right) + \sqrt{\frac{6}{7a}}\sin\left(\frac{2\pi x}{a}\right) + \frac{2}{\sqrt{7a}}\sin\left(\frac{3\pi x}{a}\right).$$

(a) If we measure energy, what values will we find and with what probabilities? Calculate the average energy.

(b) Find the wave function $\psi(x,t)$ at any later time t. Determine the probability of finding the particle at a time t in the state $\varphi(x,t) = 1/\sqrt{a}\sin(3\pi x/a)\exp(-iE_3 t/\hbar)$.

(c) Calculate the probability density $\rho(x,t)$ and the current density $\vec{J}(x,t)$. Verify that $\partial\rho/\partial t + \vec{\nabla}\cdot\vec{J}(x,t) = 0$.

Exercise 4.19

Consider a particle in an infinite square well whose wave function is given by

$$\psi(x) = \begin{cases} Ax(a^2 - x^2), & 0 < x < a, \\ 0, & \text{elsewhere,} \end{cases}$$

where A is a real constant.

(a) Find A so that $\psi(x)$ is normalized.

(b) Calculate the position and momentum uncertainties, Δx and Δp, and the product $\Delta x \Delta p$.

(c) Calculate the probability of finding $5^2\pi^2\hbar^2/(2ma^2)$ for a measurement of the energy.

Exercise 4.20

The relativistic expression for the energy of a free particle is $E^2 = m_0^2 c^4 + p^2 c^2$.

(a) Write down the corresponding relativistic Schrödinger equation, by quantizing this energy expression (i.e., replacing E and p with their corresponding operators). This equation is called the Klein–Gordon equation.

(b) Find the solutions corresponding to a free particle moving along the x-axis.

Exercise 4.21

(a) Write down the classical (gravitational) energy E_c of a particle of mass m at rest a height h_0 above the ground (take the zero potential energy to be located at the ground level).

(b) Use the uncertainty principle to estimate the ground state energy E_0 of the particle introduced in (a); note that the particle is subject to gravity. Compare E_0 to E_c.

(c) If $h_0 = 3\,\text{m}$ obtain the numerical values of E_c and the quantum mechanical correction $(E_0 - E_c)$ for a neutron and then for a particle of mass $m = 0.01\,\text{kg}$. Comment on the importance of the quantum correction in both cases.

Exercise 4.22

Find the energy levels and the wave functions of two noninteracting particles of masses m_1 and m_2 that are moving in a common infinite square well potential

$$V(x_i) = \begin{cases} 0, & 0 \le x_i \le a, \\ +\infty, & \text{elsewhere,} \end{cases}$$

where x_i is the position of the ith particle (i.e., x_i denotes $x = x_1$ or x_2).

Exercise 4.23

A particle of mass m is subject to a repulsive delta potential $V(x) = V_0\delta(x)$, where $V_0 > 0$ (V_0 has the dimensions of Energy×Distance). Find the reflection and transmission coefficients, R and T.

Exercise 4.24

A particle of mass m is scattered by a double-delta potential $V(x) = V_0\delta(x - a) + V_0\delta(x + a)$, where $V_0 > 0$.

(a) Find the transmission coefficient for the particle at an energy $E > 0$.

(b) When V_0 is very large (i.e., $V_0 \to \infty$), find the energies corresponding to the resonance case (i.e., $T = 1$) and compare them with the energies of an infinite square well potential having a width of $2a$.

Exercise 4.25

A particle of mass m is subject to an antisymmetric delta potential $V(x) = V_0\delta(x+a) - V_0\delta(x-a)$, where $V_0 > 0$.

(a) Show that there is always one and only one bound state, and find the expression that gives its energy.

(b) Find the transmission coefficient T.

Exercise 4.26

A particle of mass m is subject to a delta potential

$$V(x) = \begin{cases} \infty, & x \le 0, \\ V_0\delta(x-a), & x > 0, \end{cases}$$

where $V_0 > 0$.

(a) Find the wave functions corresponding to the cases $0 < x < a$ and $x > a$.

(b) Find the transmission coefficient.

Exercise 4.27

A particle of mass m, besides being confined to move in an infinite square well potential of size $2a$ with walls at $x = -a$ and $x = a$, is subject to an attractive delta potential

$$V(x) = \begin{cases} V_0\delta(x), & -a < x < a, \\ \infty, & \text{elsewhere}, \end{cases}$$

where $V_0 > 0$.

(a) Find the particle's wave function corresponding to *even* solutions when $E > 0$.

(b) Find the energy levels corresponding to *even* solutions.

Exercise 4.28

A particle of mass m, besides being confined to move in an infinite square well potential of size $2a$ with walls at $x = -a$ and $x = a$, is subject to an attractive delta potential

$$V(x) = \begin{cases} V_0\delta(x), & -a < x < a, \\ \infty, & \text{elsewhere}, \end{cases}$$

where $V_0 > 0$.

(a) Find the particle's wave function corresponding to *odd* solutions when $E > 0$.

(b) Find the energy levels corresponding to *odd* solutions.

Exercise 4.29

Consider a particle of mass m that is moving under the influence of an attractive delta potential

$$V(x) = \begin{cases} -V_0\delta(x), & x > -a, \\ \infty, & x < -a, \end{cases}$$

where $V_0 > 0$. Discuss the existence of bound states in terms of V_0 and a.

Exercise 4.30
Consider a system of two identical harmonic oscillators (with an angular frequency ω).

(a) Find the energy levels when the oscillators are independent (non-interacting).

(b) Find the energy levels when the oscillators are coupled by an interaction $-\lambda \hat{X}_1 \hat{X}_2$, where λ is a constant.

(c) Assuming that $\lambda \ll m\omega^2$ (weak coupling limit), find an approximate value to first order in $\lambda/m\omega^2$ for the energy expression derived in part (b).

Exercise 4.31
A particle is initially in its ground state in an infinite one-dimensional potential box with sides at $x = 0$ and $x = a$. If the wall of the box at $x = a$ is *suddenly* moved to $x = 3a$, calculate the probability of finding the particle in

(a) the ground state of the new box and

(b) the first excited state of the new box.

(c) Now, calculate the probability of finding the particle in the first excited state of the new box, assuming the particle was initially in the first excited state of the old box.

Exercise 4.32
A particle is initially in its ground state in a one-dimensional harmonic oscillator potential, $\hat{V}(x) = \frac{1}{2}k\hat{x}^2$. If the spring constant is *suddenly* doubled, calculate the probability of finding the particle in the ground state of the new potential.

Exercise 4.33
Consider an electron in an infinite potential well

$$V(x) = \begin{cases} 0, & 0 < x < a, \\ +\infty, & \text{elsewhere,} \end{cases}$$

where $a = 10^{-10}$ m.

(a) Calculate the energy levels of the three lowest states (the results should be expressed in eV) and the corresponding wavelengths of the electron.

(b) Calculate the frequency of the radiation that would cause the electron to jump from the ground to the third excited energy level.

(c) When the electron de-excites, what are the frequencies of the emitted photons?

(d) Specify the probability densities for all these three states and plot them.

Exercise 4.34
Consider an electron which is confined to move in an infinite square well of width $a = 10^{-10}m$.

(a) Find the exact energies of the 11 lowest states (express them in eV).

(b) Solve the Schrödinger equation *numerically* and find the energies of the 11 lowest states and compare them with the exact results obtained in (a). Plot the wave functions of the five lowest states.

Chapter 5

Angular Momentum

5.1 Introduction

After treating one-dimensional problems in Chapter 4, we now should deal with three-dimensional problems. However, the study of three-dimensional systems such as atoms cannot be undertaken unless we first cover the formalism of angular momentum. The current chapter, therefore, serves as an essential prelude to Chapter 6.

Angular momentum is as important in classical mechanics as in quantum mechanics. It is particularly useful for studying the dynamics of systems that move under the influence of *spherically symmetric*, or *central*, potentials, $V(\vec{r}) = V(r)$, for the orbital angular momenta of these systems are *conserved*. For instance, as mentioned in Chapter 1, one of the cornerstones of Bohr's model of the hydrogen atom (where the electron moves in the proton's Coulomb potential, a central potential) is based on the quantization of angular momentum. Additionally, angular momentum plays a critical role in the description of molecular rotations, the motion of electrons in atoms, and the motion of nucleons in nuclei. The quantum theory of angular momentum is thus a prerequisite for studying molecular, atomic, and nuclear systems.

In this chapter we are going to consider the general formalism of angular momentum. We will examine the various properties of the angular momentum operator, and then focus on determining its eigenvalues and eigenstates. Finally, we will apply this formalism to the determination of the eigenvalues and eigenvectors of the spin and orbital angular momenta.

5.2 Orbital Angular Momentum

In classical physics the angular momentum of a particle with momentum $\vec{p}$ and position $\vec{r}$ is defined by

$$\vec{L} = \vec{r} \times \vec{p} = (yp_z - zp_y)\hat{\imath} + (zp_x - xp_z)\hat{\jmath} + (xp_y - yp_x)\hat{k}. \tag{5.1}$$

The orbital angular momentum operator $\hat{\vec{L}}$ can be obtained at once by replacing $\vec{r}$ and $\vec{p}$ by the corresponding operators in the position representation, $\hat{\vec{R}}$ and $\hat{\vec{P}} = -i\hbar\vec{\nabla}$:

$$\boxed{\hat{\vec{L}} = \hat{\vec{R}} \times \hat{\vec{P}} = -i\hbar\hat{\vec{R}} \times \vec{\nabla}\,.} \tag{5.2}$$

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

The Cartesian components of $\hat{\vec{L}}$ are

$$\hat{L}_x = \hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y = -i\hbar\left(\hat{Y}\frac{\partial}{\partial z} - \hat{Z}\frac{\partial}{\partial y}\right), \tag{5.3}$$

$$\hat{L}_y = \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z = -i\hbar\left(\hat{Z}\frac{\partial}{\partial x} - \hat{X}\frac{\partial}{\partial z}\right), \tag{5.4}$$

$$\hat{L}_z = \hat{X}\hat{P}_y - \hat{Y}\hat{P}_x = -i\hbar\left(\hat{X}\frac{\partial}{\partial y} - \hat{Y}\frac{\partial}{\partial x}\right). \tag{5.5}$$

Clearly, angular momentum does not exist in a one-dimensional space. We should mention that the components $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$, and the square of $\hat{\vec{L}}$,

$$\hat{\vec{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \tag{5.6}$$

are all Hermitian.

Commutation relations

Since $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ mutually commute and so do $\hat{P}_x$, $\hat{P}_y$, and $\hat{P}_z$, and since $[\hat{X}, \hat{P}_x] = i\hbar$, $[\hat{Y}, \hat{P}_y] = i\hbar$, $[\hat{Z}, \hat{P}_z] = i\hbar$, we have

$$\begin{aligned}
[\hat{L}_x, \hat{L}_y] &= [\hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y,\ \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z] \\
&= [\hat{Y}\hat{P}_z, \hat{Z}\hat{P}_x] - [\hat{Y}\hat{P}_z, \hat{X}\hat{P}_z] - [\hat{Z}\hat{P}_y, \hat{Z}\hat{P}_x] + [\hat{Z}\hat{P}_y, \hat{X}\hat{P}_z] \\
&= \hat{Y}[\hat{P}_z, \hat{Z}]\hat{P}_x + \hat{X}[\hat{Z}, \hat{P}_z]\hat{P}_y = i\hbar(\hat{X}\hat{P}_y - \hat{Y}\hat{P}_x) \\
&= i\hbar\hat{L}_z.
\end{aligned} \tag{5.7}$$

A similar calculation yields the other two commutation relations; but it is much simpler to infer them from (5.7) by means of a *cyclic permutation* of the xyz components, $x \to y \to z \to x$:

$$\boxed{[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \qquad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \qquad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y.} \tag{5.8}$$

As mentioned in Chapter 3, since $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ do not commute, we cannot measure them simultaneously to arbitrary accuracy.

Note that the commutation relations (5.8) were derived by expressing the orbital angular momentum in the *position representation*, but since these are operator relations, they must be valid in any representation. In the following section we are going to consider the general formalism of angular momentum, a formalism that is restricted to no particular representation.

Example 5.1

(a) Calculate the commutators $[\hat{X}, \hat{L}_x]$, $[\hat{X}, \hat{L}_y]$, and $[\hat{X}, \hat{L}_z]$.

(b) Calculate the commutators: $[\hat{P}_x, \hat{L}_x]$, $[\hat{P}_x, \hat{L}_y]$, and $[\hat{P}_x, \hat{L}_z]$.

(c) Use the results of (a) and (b) to calculate $[\hat{X}, \hat{\vec{L}}^2]$ and $[\hat{P}_x, \hat{\vec{L}}^2]$.

Solution

(a) The only nonzero commutator which involves $\hat{X}$ and the various components of $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$ is $[\hat{X}, \hat{P}_x] = i\hbar$. Having stated this result, we can easily evaluate the needed commutators. First, since $\hat{L}_x = \hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y$ involves no $\hat{P}_x$, the operator $\hat{X}$ commutes separately with $\hat{Y}$, $\hat{P}_z$, $\hat{Z}$, and $\hat{P}_y$; hence

$$[\hat{X}, \hat{L}_x] = [\hat{X}, \hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y] = 0. \tag{5.9}$$

The evaluation of the other two commutators is straightforward:

$$[\hat{X}, \hat{L}_y] = [\hat{X}, \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z] = [\hat{X}, \hat{Z}\hat{P}_x] = \hat{Z}[\hat{X}, \hat{P}_x] = i\hbar\hat{Z}, \tag{5.10}$$

$$[\hat{X}, \hat{L}_z] = [\hat{X}, \hat{X}\hat{P}_y - \hat{Y}\hat{P}_x] = -[\hat{X}, \hat{Y}\hat{P}_x] = -\hat{Y}[\hat{X}, \hat{P}_x] = -i\hbar\hat{Y}. \tag{5.11}$$

(b) The only commutator between $\hat{P}_x$ and the components of $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$ that survives is again $[\hat{P}_x, \hat{X}] = -i\hbar$. We may thus infer

$$[\hat{P}_x, \hat{L}_x] = [\hat{P}_x, \hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y] = 0, \tag{5.12}$$

$$[\hat{P}_x, \hat{L}_y] = [\hat{P}_x, \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z] = -[\hat{P}_x, \hat{X}\hat{P}_z] = -[\hat{P}_x, \hat{X}]\hat{P}_z = i\hbar\hat{P}_z, \tag{5.13}$$

$$[\hat{P}_x, \hat{L}_z] = [\hat{P}_x, \hat{X}\hat{P}_y - \hat{Y}\hat{P}_x] = [\hat{P}_x, \hat{X}\hat{P}_y] = [\hat{P}_x, \hat{X}]\hat{P}_y = -i\hbar\hat{P}_y. \tag{5.14}$$

(c) Using the commutators derived in (a) and (b), we infer

$$\begin{aligned}[\hat{X}, \hat{\vec{L}}^2] &= [\hat{X}, \hat{L}_x^2] + [\hat{X}, \hat{L}_y^2] + [\hat{X}, \hat{L}_z^2] \\ &= 0 + \hat{L}_y[\hat{X}, \hat{L}_y] + [\hat{X}, \hat{L}_y]\hat{L}_y + \hat{L}_z[\hat{X}, \hat{L}_z] + [\hat{X}, \hat{L}_z]\hat{L}_z \\ &= i\hbar(\hat{L}_y\hat{Z} + \hat{Z}\hat{L}_y - \hat{L}_z\hat{Y} - \hat{Y}\hat{L}_y),\end{aligned} \tag{5.15}$$

$$\begin{aligned}[\hat{P}_x, \hat{\vec{L}}^2] &= [\hat{P}_x, \hat{L}_x^2] + [\hat{P}_x, \hat{L}_y^2] + [\hat{P}_x, \hat{L}_z^2] \\ &= 0 + \hat{L}_y[\hat{P}_x, \hat{L}_y] + [\hat{P}_x, \hat{L}_y]\hat{L}_y + \hat{L}_z[\hat{P}_x, \hat{L}_z] + [\hat{P}_x, \hat{L}_z]\hat{L}_z \\ &= i\hbar(\hat{L}_y\hat{P}_z + \hat{P}_z\hat{L}_y - \hat{L}_z\hat{P}_y - \hat{P}_y\hat{L}_y).\end{aligned} \tag{5.16}$$

5.3 General Formalism of Angular Momentum

Let us now introduce a more general angular momentum operator $\hat{\vec{J}}$ that is defined by its three components $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$, which satisfy the following commutation relations:

$$\boxed{[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z, \qquad [\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x, \qquad [\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y,} \tag{5.17}$$

or equivalently by

$$\hat{\vec{J}} \times \hat{\vec{J}} = i\hbar\hat{\vec{J}}. \tag{5.18}$$

Since $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$ do not mutually commute, they cannot be simultaneously diagonalized; that is, they do not possess common eigenstates. The square of the angular momentum,

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2, \tag{5.19}$$

is a scalar operator; hence it commutes with $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$:

$$[\hat{\vec{J}}^2, \hat{J}_k] = 0, \tag{5.20}$$

where k stands for x, y, and z. For instance, in the the case $k = x$ we have

$$\begin{aligned}[\hat{\vec{J}}^2, \hat{J}_x] &= [\hat{J}_x^2, \hat{J}_x] + \hat{J}_y[\hat{J}_y, \hat{J}_x] + [\hat{J}_y, \hat{J}_x]\hat{J}_y + \hat{J}_z[\hat{J}_z, \hat{J}_x] + [\hat{J}_z, \hat{J}_x]\hat{J}_z \\ &= \hat{J}_y(-i\hbar\hat{J}_z) + (-i\hbar\hat{J}_z)\hat{J}_y + \hat{J}_z(i\hbar\hat{J}_y) + (i\hbar\hat{J}_y)\hat{J}_z \\ &= 0, \end{aligned} \tag{5.21}$$

because $[\hat{J}_x^2, \hat{J}_x] = 0$, $[\hat{J}_y, \hat{J}_x] = -i\hbar\hat{J}_z$, and $[\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y$. We should note that the operators $\hat{J}_x$, $\hat{J}_y$, $\hat{J}_z$, and $\hat{\vec{J}}^2$ are all Hermitian; their eigenvalues are real.

Eigenstates and eigenvalues of the angular momentum operator

Since $\hat{\vec{J}}^2$ commutes with $\hat{J}_x$, $\hat{J}_y$ and $\hat{J}_z$, each component of $\hat{J}$ can be separately diagonalized (hence it has simultaneous eigenfunctions) with $\hat{\vec{J}}^2$. But since the components $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$ do not mutually commute, we can choose only one of them to be simultaneously diagonalized with $\hat{\vec{J}}^2$. By convention we choose $\hat{J}_z$. There is nothing special about the z-direction; we can just as well take $\hat{\vec{J}}^2$ and $\hat{J}_x$ or $\hat{\vec{J}}^2$ and $\hat{J}_y$.

Let us now look for the joint eigenstates of $\hat{\vec{J}}^2$ and $\hat{J}_z$ and their corresponding eigenvalues. Denoting the joint eigenstates by $|\alpha, \beta\rangle$ and the eigenvalues of $\hat{\vec{J}}^2$ and $\hat{J}_z$ by $\hbar^2\alpha$ and $\hbar\beta$, respectively, we have

$$\hat{\vec{J}}^2 |\alpha, \beta\rangle = \hbar^2\alpha |\alpha, \beta\rangle, \tag{5.22}$$

$$\hat{J}_z |\alpha, \beta\rangle = \hbar\beta |\alpha, \beta\rangle. \tag{5.23}$$

The factor $\hbar$ is introduced so that α and β are dimensionless; recall that the angular momentum has the dimensions of $\hbar$ and that the physical dimensions of $\hbar$ are: $[\hbar]$ = energy × time. For simplicity, we will assume that these eigenstates are orthonormal:

$$\langle\alpha', \beta' | \alpha, \beta\rangle = \delta_{\alpha',\alpha}\delta_{\beta',\beta}\,. \tag{5.24}$$

Now we need to introduce *raising* and *lowering* operators $\hat{J}_+$ and $\hat{J}_-$, just as we did when we studied the harmonic oscillator in Chapter 4:

$$\boxed{\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y.} \tag{5.25}$$

This leads to

$$\hat{J}_x = \frac{1}{2}(\hat{J}_+ + \hat{J}_-), \qquad \hat{J}_y = \frac{1}{2i}(\hat{J}_+ - \hat{J}_-); \tag{5.26}$$

hence

$$\hat{J}_x^2 = \frac{1}{4}(\hat{J}_+^2 + \hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+ + \hat{J}_-^2), \qquad \hat{J}_y^2 = -\frac{1}{4}(\hat{J}_+^2 - \hat{J}_+\hat{J}_- - \hat{J}_-\hat{J}_+ + \hat{J}_-^2). \tag{5.27}$$

Using (5.17) we can easily obtain the following commutation relations:

$$[\hat{\vec{J}}^2, \hat{J}_\pm] = 0, \qquad [\hat{J}_+, \hat{J}_-] = 2\hbar\hat{J}_z, \qquad [\hat{J}_z, \hat{J}_\pm] = \pm\hbar\hat{J}_\pm. \tag{5.28}$$

In addition, $\hat{J}_+$ and $\hat{J}_-$ satisfy

$$\hat{J}_+\hat{J}_- = \hat{J}_x^2 + \hat{J}_y^2 + \hbar\hat{J}_z = \hat{\vec{J}}^2 - \hat{J}_z^2 + \hbar\hat{J}_z, \tag{5.29}$$

$$\hat{J}_-\hat{J}_+ = \hat{J}_x^2 + \hat{J}_y^2 - \hbar\hat{J}_z = \hat{\vec{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z. \tag{5.30}$$

These relations lead to

$$\boxed{\hat{\vec{J}}^2 = \hat{J}_\pm\hat{J}_\mp + \hat{J}_z^2 \mp \hbar\hat{J}_z,} \tag{5.31}$$

which in turn yield

$$\boxed{\hat{\vec{J}}^2 = \frac{1}{2}(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+) + \hat{J}_z^2.} \tag{5.32}$$

Let us see how $\hat{J}_\pm$ operate on $|\alpha, \beta\rangle$. First, since $\hat{J}_\pm$ do not commute with $\hat{J}_z$, the kets $|\alpha, \beta\rangle$ are not eigenstates of $\hat{J}_\pm$. Using the relations (5.28) we have

$$\hat{J}_z(\hat{J}_\pm|\alpha, \beta\rangle) = (\hat{J}_\pm\hat{J}_z \pm \hbar\hat{J}_\pm)|\alpha, \beta\rangle = \hbar(\beta \pm 1)(\hat{J}_\pm|\alpha, \beta\rangle); \tag{5.33}$$

hence the ket $(\hat{J}_\pm|\alpha, \beta\rangle)$ is an eigenstate of $\hat{J}_z$ with eigenvalues $\hbar(\beta \pm 1)$. Now since $\hat{J}_z$ and $\hat{\vec{J}}^2$ commute, $(\hat{J}_\pm|\alpha, \beta\rangle)$ must also be an eigenstate of $\hat{\vec{J}}^2$. The eigenvalue of $\hat{\vec{J}}^2$ when acting on $\hat{J}_\pm|\alpha, \beta\rangle$ can be determined by making use of the commutator $[\hat{\vec{J}}^2, \hat{J}_\pm] = 0$. The state $(\hat{J}_\pm|\alpha, \beta\rangle)$ is also an eigenstate of $\hat{\vec{J}}^2$ with eigenvalue $\hbar^2\alpha$:

$$\hat{\vec{J}}^2(\hat{J}_\pm|\alpha, \beta\rangle) = \hat{J}_\pm\hat{\vec{J}}^2|\alpha, \beta\rangle = \hbar^2\alpha(\hat{J}_\pm|\alpha, \beta\rangle). \tag{5.34}$$

From (5.33) and (5.34) we infer that when $\hat{J}_\pm$ acts on $|\alpha, \beta\rangle$, it does not affect the first quantum number α, but it raises or lowers the second quantum number β by one unit. That is, $\hat{J}_\pm|\alpha, \beta\rangle$ is proportional to $|\alpha, \beta \pm 1\rangle$:

$$\hat{J}_\pm|\alpha, \beta\rangle = C^\pm_{\alpha\beta}|\alpha, \beta \pm 1\rangle. \tag{5.35}$$

We will determine the constant $C^\pm_{\alpha\beta}$ later on.

Note that, for a given eigenvalue α of $\hat{\vec{J}}^2$, there exists an *upper limit* for the quantum number β. This is due to the fact that the operator $\hat{\vec{J}}^2 - \hat{J}_z^2$ is positive, since the matrix elements of $\hat{\vec{J}}^2 - \hat{J}_z^2 = \hat{J}_x^2 + \hat{J}_y^2$ are ≥ 0; we can therefore write

$$\langle\alpha, \beta|\hat{\vec{J}}^2 - \hat{J}_z^2|\alpha, \beta\rangle = \hbar^2(\alpha - \beta^2) \geq 0, \quad \Longrightarrow \quad \alpha \geq \beta^2. \tag{5.36}$$

Since β has an upper limit β_{max}, there must exist a state $|\alpha, \beta_{max}\rangle$ which cannot be raised further:

$$\hat{J}_+|\alpha, \beta_{max}\rangle = 0. \tag{5.37}$$

Using this relation along with $\hat{J}_-\hat{J}_+ = \hat{\vec{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z$, we see that $\hat{J}_-\hat{J}_+|\alpha, \beta_{max}\rangle = 0$ or

$$(\hat{\vec{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)|\alpha, \beta_{max}\rangle = \hbar^2(\alpha - \beta_{max}^2 - \beta_{max})|\alpha, \beta_{max}\rangle; \tag{5.38}$$

hence

$$\alpha = \beta_{max}(\beta_{max} + 1). \tag{5.39}$$

After n successive applications of $\hat{J}_-$ on $|\ \alpha,\ \beta_{max}\rangle$, we must be able to reach a state $|\ \alpha,\ \beta_{min}\rangle$ which cannot be lowered further:

$$\hat{J}_-\ |\ \alpha,\ \beta_{min}\rangle = 0. \tag{5.40}$$

Using $\hat{J}_-\hat{J}_- = \hat{\vec{J}}^2 - \hat{J}_z^2 + \hbar\hat{J}_z$, and by analogy with (5.38) and (5.39), we infer that

$$\alpha = \beta_{min}(\beta_{min} - 1). \tag{5.41}$$

Comparing (5.39) and (5.41) we obtain

$$\beta_{max} = -\beta_{min}. \tag{5.42}$$

Since β_{min} was reached by n applications of $\hat{J}_-$ on $|\ \alpha,\ \beta_{max}\rangle$, it follows that

$$\beta_{max} = \beta_{min} + n, \tag{5.43}$$

and since $\beta_{min} = -\beta_{max}$ we conclude that

$$\beta_{max} = \frac{n}{2}. \tag{5.44}$$

Hence β_{max} can be integer or half-odd-integer, depending on n being even or odd.

It is now appropriate to introduce the notation j and m to denote β_{max} and β, respectively:

$$j = \beta_{max} = \frac{n}{2}, \qquad m = \beta; \tag{5.45}$$

hence the eigenvalue of $\hat{\vec{J}}^2$ is given by

$$\alpha = j(j + 1). \tag{5.46}$$

Now since $\beta_{min} = -\beta_{max}$, and with n positive, we infer that the allowed values of m lie between $-j$ and $+j$:

$$\boxed{-j \le m \le j.} \tag{5.47}$$

The results obtained thus far can be summarized as follows: the eigenvalues of $\hat{\vec{J}}^2$ and J_z corresponding to the joint eigenvectors $|\ j,\ m\rangle$ are given, respectively, by $\hbar^2 j(j + 1)$ and $\hbar m$:

$$\boxed{\hat{\vec{J}}^2\ |\ j,\ m\rangle = \hbar^2 j(j + 1)\ |\ j,\ m\rangle \quad \text{and} \quad \hat{J}_z\ |\ j,\ m\rangle = \hbar m\ |\ j,\ m\rangle,} \tag{5.48}$$

where $j = 0, 1/2, 1, 3/2, \ldots$ and $m = -j, -(j-1), \ldots, j-1, j$. So for each j there are $2j + 1$ values of m. For example, if $j = 1$ then m takes the three values $-1, 0, 1$; and if $j = 5/2$ then m takes the six values $-5/2, -3/2, -1/2, 1/2, 3/2, 5/2$. The values of j are either integer or half-integer. We see that the spectra of the angular momentum operators $\hat{\vec{J}}^2$ and $\hat{J}_z$ are discrete. Since the eigenstates corresponding to different angular momenta are orthogonal, and since the angular momentum spectra are discrete, the orthonormality condition is

$$\boxed{\langle j',\ m'\ |\ j,\ m\rangle = \delta_{j',j}\delta_{m',m}.} \tag{5.49}$$

Let us now determine the eigenvalues of $\hat{J}_\pm$ within the $\{|\ j,\ m\rangle\}$ basis; $|\ j,\ m\rangle$ is not an eigenstate of $\hat{J}_\pm$. We can rewrite equation (5.35) as

$$\hat{J}_\pm\ |\ j,\ m\rangle = C^\pm_{jm}\ |\ j,\ m \pm 1\rangle. \tag{5.50}$$

We are going to derive C^+_{jm} and then infer C^-_{jm}. Since $|\ j,\ m\rangle$ is normalized, we can use (5.50) to obtain the following two expressions:

$$(\hat{J}_+\ |\ j,\ m\rangle)^\dagger\ (\hat{J}_+\ |\ j,\ m\rangle) = |C^+_{jm}|^2\langle j,\ m+1\ |\ j,\ m+1\rangle = |C^+_{jm}|^2, \tag{5.51}$$

$$\left|C^+_{jm}\right|^2 = \langle j,\ m|\hat{J}_-\hat{J}_+\ |\ j,\ m\rangle. \tag{5.52}$$

But since $\hat{J}_-\hat{J}_+$ is equal to $(\hat{\vec{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)$, and assuming the arbitrary phase of C^+_{jm} to be zero, we conclude that

$$C^+_{jm} = \sqrt{\langle j,\ m|\hat{\vec{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z\ |\ j,\ m\rangle} = \hbar\sqrt{j(j+1) - m(m+1)}. \tag{5.53}$$

By analogy with C^+_{jm} we can easily infer the expression for C^-_{jm}:

$$C^-_{jm} = \hbar\sqrt{j(j+1) - m(m-1)}. \tag{5.54}$$

Thus, the equations for $\hat{J}_+$ and $\hat{J}_-$ are given by

$$\boxed{\hat{J}_\pm\ |\ j,\ m\rangle = \hbar\sqrt{j(j+1) - m(m\pm 1)}\ |\ j,\ m\pm 1\rangle} \tag{5.55}$$

or

$$\boxed{\hat{J}_\pm\ |\ j,\ m\rangle = \hbar\sqrt{(j\mp m)(j\pm m+1)}\ |\ j,\ m\pm 1\rangle,} \tag{5.56}$$

which in turn leads to the two following relations:

$$\begin{aligned}\hat{J}_x\ |\ j,\ m\rangle &= \frac{1}{2}(\hat{J}_+ + \hat{J}_-)\ |\ j,\ m\rangle \\ &= \frac{\hbar}{2}\left[\sqrt{(j-m)(j+m+1)}\ |\ j,\ m+1\rangle + \sqrt{(j+m)(j-m+1)}\ |\ j,\ m-1\rangle\right],\end{aligned} \tag{5.57}$$

$$\begin{aligned}\hat{J}_y\ |\ j,\ m\rangle &= \frac{1}{2i}(\hat{J}_+ - \hat{J}_-)\ |\ j,\ m\rangle \\ &= \frac{\hbar}{2i}\left[\sqrt{(j-m)(j+m+1)}\ |\ j,\ m+1\rangle - \sqrt{(j+m)(j-m+1)}\ |\ j,\ m-1\rangle\right].\end{aligned} \tag{5.58}$$

The expectation values of $\hat{J}_x$ and $\hat{J}_y$ are therefore zero:

$$\langle j,\ m\ |\ \hat{J}_x\ |\ j,\ m\rangle = \langle j,\ m\ |\ \hat{J}_y\ |\ j,\ m\rangle = 0 \tag{5.59}$$

We will show later in (5.217) that the expectation values $\langle j,\ m\ |\ \hat{J}_x^2\ |\ j,\ m\rangle$ and $\langle j,\ m\ |\ \hat{J}_y^2\ |\ j,\ m\rangle$ are equal and given by

$$\boxed{\langle \hat{J}_x^2\rangle = \langle \hat{J}_y^2\rangle = \frac{1}{2}\left[\langle j,\ m\ |\ \hat{\vec{J}}^2\ |\ j,\ m\rangle - \langle j,\ m\ |\ \hat{J}_z^2\ |\ j,\ m\rangle\right] = \frac{\hbar^2}{2}\left[j(j+1) - m^2\right].} \tag{5.60}$$

Example 5.2
Calculate $[\hat{J}_x^2, \hat{J}_y]$, $[\hat{J}_z^2, \hat{J}_y]$, and $[\hat{\vec{J}}^2, \hat{J}_y]$; then show $\langle j, m | \hat{J}_x^2 | j, m\rangle = \langle j, m | \hat{J}_y^2 | j, m\rangle$.

Solution
Since $[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$ and $[\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y$, we have

$$[\hat{J}_x^2, \hat{J}_y] = \hat{J}_x[\hat{J}_x, \hat{J}_y] + [\hat{J}_x, \hat{J}_y]\hat{J}_x = i\hbar(\hat{J}_x\hat{J}_z + \hat{J}_z\hat{J}_x) = i\hbar(2\hat{J}_x\hat{J}_z + i\hbar\hat{J}_y). \tag{5.61}$$

Similarly, since $[\hat{J}_z, \hat{J}_y] = -i\hbar\hat{J}_x$ and $[\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y$, we have

$$[\hat{J}_z^2, \hat{J}_y] = \hat{J}_z[\hat{J}_z, \hat{J}_y] + [\hat{J}_z, \hat{J}_y]\hat{J}_z = -i\hbar(\hat{J}_z\hat{J}_x + \hat{J}_x\hat{J}_z) = -i\hbar(2\hat{J}_x\hat{J}_z + i\hbar\hat{J}_y). \tag{5.62}$$

The previous two expressions yield

$$\begin{aligned}[\hat{\vec{J}}^2, \hat{J}_y] &= [\hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2, \hat{J}_y] = [\hat{J}_x^2, \hat{J}_y] + [\hat{J}_z^2, \hat{J}_y] \\ &= i\hbar(2\hat{J}_x\hat{J}_z + i\hbar\hat{J}_y) - i\hbar(2\hat{J}_x\hat{J}_z + i\hbar\hat{J}_y) = 0.\end{aligned} \tag{5.63}$$

Since we have

$$\hat{J}_x^2 = \frac{1}{4}(\hat{J}_+^2 + \hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+ + \hat{J}_-^2), \qquad \hat{J}_y^2 = -\frac{1}{4}(\hat{J}_+^2 - \hat{J}_+\hat{J}_- - \hat{J}_-\hat{J}_+ + \hat{J}_-^2), \tag{5.64}$$

and since $\langle j, m | \hat{J}_+^2 | j, m\rangle = \langle j, m | \hat{J}_-^2 | j, m\rangle = 0$, we can write

$$\langle j, m | \hat{J}_x^2 | j, m\rangle = \frac{1}{4}\langle j, m | \hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+ | j, m\rangle = \langle j, m | \hat{J}_y^2 | j, m\rangle. \tag{5.65}$$

5.4 Matrix Representation of Angular Momentum

The formalism of the previous section is general and independent of any particular representation. There are many ways to represent the angular momentum operators and their eigenstates. In this section we are going to discuss the matrix representation of angular momentum where eigenkets and operators will be represented by column vectors and square matrices, respectively. This is achieved by expanding states and operators in a discrete basis. We will see later how to represent the orbital angular momentum in the position representation.

Since $\hat{\vec{J}}^2$ and $\hat{J}_z$ commute, the set of their common eigenstates $\{| j, m\rangle\}$ can be chosen as a basis; this basis is discrete, orthonormal, and complete. For a given value of j, the orthonormalization condition for this base is given by (5.49), and the completeness condition is expressed by

$$\sum_{m=-j}^{+j} | j, m\rangle\langle j, m | = \hat{I}, \tag{5.66}$$

where $\hat{I}$ is the unit matrix. The operators $\hat{J}^2$ and $\hat{J}_z$ are diagonal in the basis given by their joint eigenstates

$$\langle j', m' \mid \hat{J}^2 \mid j, m\rangle = \hbar^2 j(j+1)\delta_{j',j}\delta_{m',m} \tag{5.67}$$

$$\langle j', m' \mid \hat{J}_z \mid j, m\rangle = \hbar m \delta_{j',j}\delta_{m',m}. \tag{5.68}$$

Thus, the matrices representing $\hat{J}^2$ and $\hat{J}_z$ in the $\{|\, j, m\rangle\}$ eigenbasis are diagonal, their diagonal elements being equal to $\hbar^2 j(j+1)$ and $\hbar m$, respectively.

Now since the operators $\hat{J}_\pm$ do not commute with $\hat{J}_z$, they are represented in the $\{|\, j, m\rangle\}$ basis by matrices that are not diagonal:

$$\boxed{\langle j', m' \mid \hat{J}_\pm \mid j, m\rangle = \hbar\sqrt{j(j+1) - m(m\pm 1)}\,\delta_{j',j}\delta_{m',m\pm 1}.} \tag{5.69}$$

We can infer the matrices of $\hat{J}_x$ and $\hat{J}_y$ from (5.57) and (5.58):

$$\begin{aligned}\langle j', m' \mid \hat{J}_x \mid j, m\rangle = \frac{\hbar}{2}\Big[&\sqrt{j(j+1) - m(m+1)}\,\delta_{m',m+1} \\ &+ \sqrt{j(j+1) - m(m-1)}\,\delta_{m',m-1}\Big]\delta_{j',j},\end{aligned} \tag{5.70}$$

$$\begin{aligned}\langle j', m' \mid \hat{J}_y \mid j, m\rangle = \frac{\hbar}{2i}\Big[&\sqrt{j(j+1) - m(m+1)}\,\delta_{m',m+1} \\ &- \sqrt{j(j+1) - m(m-1)}\,\delta_{m',m-1}\Big]\delta_{j',j}.\end{aligned} \tag{5.71}$$

Example 5.3 (Angular momentum $j = 1$)
Consider the case where $j = 1$.

(a) Find the matrices representing the operators $\hat{J}^2$, $\hat{J}_z$, $\hat{J}_\pm$, $\hat{J}_x$, and $\hat{J}_y$.

(b) Find the joint eigenstates of $\hat{J}^2$ and $\hat{J}_z$ and verify that they form an orthonormal and complete basis.

(c) Use the matrices of $\hat{J}_x$, $\hat{J}_y$ and $\hat{J}_z$ to calculate $[\hat{J}_x, \hat{J}_y]$, $[\hat{J}_y, \hat{J}_z]$, and $[\hat{J}_z, \hat{J}_x]$.

(d) Verify that $\hat{J}_z^3 = \hbar^2\hat{J}_z$ and $\hat{J}_\pm^3 = 0$.

Solution

(a) For $j = 1$ the allowed values of m are -1, 0, 1. The joint eigenstates of $\hat{J}^2$ and $\hat{J}_z$ are $|\,1, -1\rangle$, $|\,1, 0\rangle$, and $|\,1, 1\rangle$. The matrix representations of the operators $\hat{J}^2$ and $\hat{J}_z$ can be inferred

from (5.67) and (5.68):

$$\hat{J}^2 = \begin{pmatrix} \langle 1,1|\hat{J}^2|1,1\rangle & \langle 1,1|\hat{J}^2|1,0\rangle & \langle 1,1|\hat{J}^2|1,-1\rangle \\ \langle 1,0|\hat{J}^2|1,1\rangle & \langle 1,0|\hat{J}^2|1,0\rangle & \langle 1,0|\hat{J}^2|1,-1\rangle \\ \langle 1,-1|\hat{J}^2|1,1\rangle & \langle 1,-1|\hat{J}^2|1,0\rangle & \langle 1,-1|\hat{J}^2|1,-1\rangle \end{pmatrix}$$

$$= 2\hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{5.72}$$

$$\hat{J}_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{5.73}$$

Similarly, using (5.69), we can ascertain that the matrices of $\hat{J}_+$ and $\hat{J}_-$ are given by

$$\hat{J}_- = \hbar\sqrt{2} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \hat{J}_+ = \hbar\sqrt{2} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.74}$$

The matrices for $\hat{J}_x$ and $\hat{J}_y$ in the $\{|\, j,\, m\rangle\}$ basis result immediately from the relations $\hat{J}_x = (\hat{J}_+ + \hat{J}_-)/2$ and $\hat{J}_y = i(\hat{J}_- - \hat{J}_+)/2$:

$$\hat{J}_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \hat{J}_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{5.75}$$

(b) The joint eigenvectors of $\hat{J}^2$ and $\hat{J}_z$ can be obtained as follows. The matrix equation of $\hat{J}_z\,|\, j,\, m\rangle = m\hbar\,|\, j,\, m\rangle$ is

$$\hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = m\hbar \begin{pmatrix} a \\ b \\ c \end{pmatrix} \implies \begin{matrix} \hbar a = m\hbar a \\ 0 = m\hbar b \\ -\hbar c = m\hbar c \end{matrix}. \tag{5.76}$$

The normalized solutions to these equations for $m = 1, 0, -1$ are respectively given by $a = 1$, $b = c = 0$; $a = 0$, $b = 1$, $c = 0$; and $a = b = 0$, $c = 1$; that is,

$$|\,1,\,1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |\,1,\,0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |\,1,\,-1\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{5.77}$$

We can verify that these vectors are orthonormal:

$$\langle 1,\, m'\,|\,1,\, m\rangle = \delta_{m',m} \qquad (m', m = -1, 0, 1). \tag{5.78}$$

We can also verify that they are complete:

$$\sum_{m=-1}^{1} |\,1,\, m\rangle\langle 1,\, m| = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}(0\,0\,1) + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}(0\,1\,0) + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}(1\,0\,0)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.79}$$

(c) Using the matrices (5.75) we have

$$\hat{J}_x\hat{J}_y = \frac{\hbar^2}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} = \frac{\hbar^2}{2}\begin{pmatrix} i & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & -i \end{pmatrix}, \tag{5.80}$$

$$\hat{J}_y\hat{J}_x = \frac{\hbar^2}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \frac{\hbar^2}{2}\begin{pmatrix} -i & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & i \end{pmatrix}; \tag{5.81}$$

hence

$$\hat{J}_x\hat{J}_y - \hat{J}_y\hat{J}_x = \frac{\hbar^2}{2}\begin{pmatrix} 2i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2i \end{pmatrix} = i\hbar^2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = i\hbar\hat{J}_z, \tag{5.82}$$

where the matrix of $\hat{J}_z$ is given by (5.73). A similar calculation leads to $[\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x$ and $[\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y$.

(d) The calculation of $\hat{J}_z^3$ and $\hat{J}_\pm^3$ is straightforward:

$$\hat{J}_z^3 = \hbar^3\left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}\right]^3 = \hbar^3\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \hbar^2\hat{J}_z, \tag{5.83}$$

$$\hat{J}_+^3 = 2\hbar^3\sqrt{2}\left[\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\right]^3 = 2\hbar^3\sqrt{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0, \tag{5.84}$$

and

$$\hat{J}_-^3 = 2\hbar^3\sqrt{2}\left[\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\right]^3 = 2\hbar^3\sqrt{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0. \tag{5.85}$$

5.5 Geometrical Representation of Angular Momentum

At issue here is the relationship between the angular momentum and its z-component; this relation can be represented geometrically as follows. For a fixed value of j, the total angular momentum $\hat{J}$ may be represented by a vector whose length, as displayed in Figure 5.1, is given by $\sqrt{\langle\hat{J}^2\rangle} = \hbar\sqrt{j(j+1)}$ and whose z-component is $\langle\hat{J}_z\rangle = \hbar m$. Since $\hat{J}_x$ and $\hat{J}_y$ are separately undefined, only their sum $\hat{J}_x^2 + \hat{J}_y^2 = \hat{J}^2 - \hat{J}_z^2$, which lies within the xy plane, is well defined. In classical terms, we can think of $\hat{J}$ as representable graphically by a vector, whose endpoint lies on a circle of radius $\hbar\sqrt{j(j+1)}$, rotating along the surface of a cone of half-angle

$$\theta = \cos^{-1}\left(\frac{m}{\sqrt{j(j+1)}}\right), \tag{5.86}$$

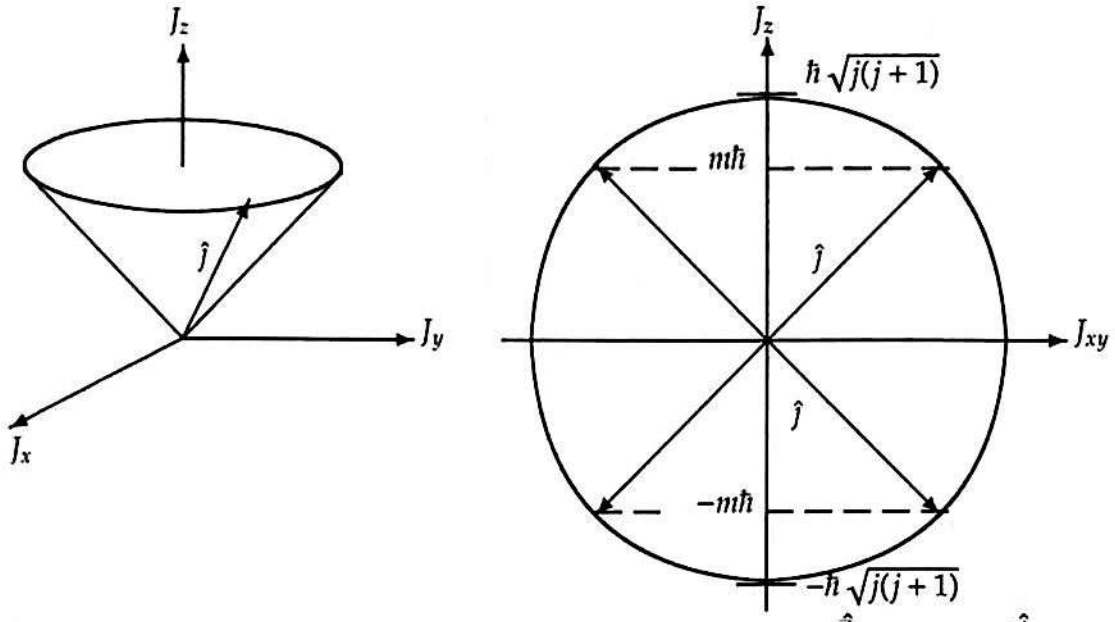


Figure 5.1: Geometrical representation of the angular momentum $\hat{\vec{J}}$: the vector $\hat{\vec{J}}$ rotates along the surface of a cone about its axis; the cone's height is equal to $m\hbar$, the projection of $\hat{\vec{J}}$ on the cone's axis. The tip of $\hat{\vec{J}}$ lies, within the J_zJ_{xy} plane, on a circle of radius $\hbar\sqrt{j(j+1)}$.

such that its projection along the z-axis is always $m\hbar$. Notice that, as the values of the quantum number m are limited to $m = -j, -j+1, \ldots, j-1, j$, the angle θ is quantized; the only possible values of θ consist of a discrete set of $2j+1$ values:

$$\theta = \cos^{-1}\left(\frac{-j}{\sqrt{j(j+1)}}\right), \cos^{-1}\left(\frac{-j+1}{\sqrt{j(j+1)}}\right), \ldots, \cos^{-1}\left(\frac{j-1}{\sqrt{j(j+1)}}\right), \cos^{-1}\left(\frac{j}{\sqrt{j(j+1)}}\right). \tag{5.87}$$

Since all orientations of $\hat{\vec{J}}$ on the surface of the cone are equally likely, the projection of $\hat{\vec{J}}$ on both the x and y axes average out to zero:

$$\langle \hat{J}_x \rangle = \langle \hat{J}_y \rangle = 0, \tag{5.88}$$

where $\langle \hat{J}_x \rangle$ stands for $\langle j, m \mid \hat{J}_x \mid j, m \rangle$.

As an example, Figure 5.2 shows the graphical representation for the $j = 2$ case. As specified in (5.87), θ takes only a discrete set of values. In this case where $j = 2$, the angle θ takes only five values corresponding respectively to $m = -2, -1, 0, 1, 2$; they are given by

$$\theta = -35.26°, -65.91°, 90°, 65.91°, 35.26°. \tag{5.89}$$

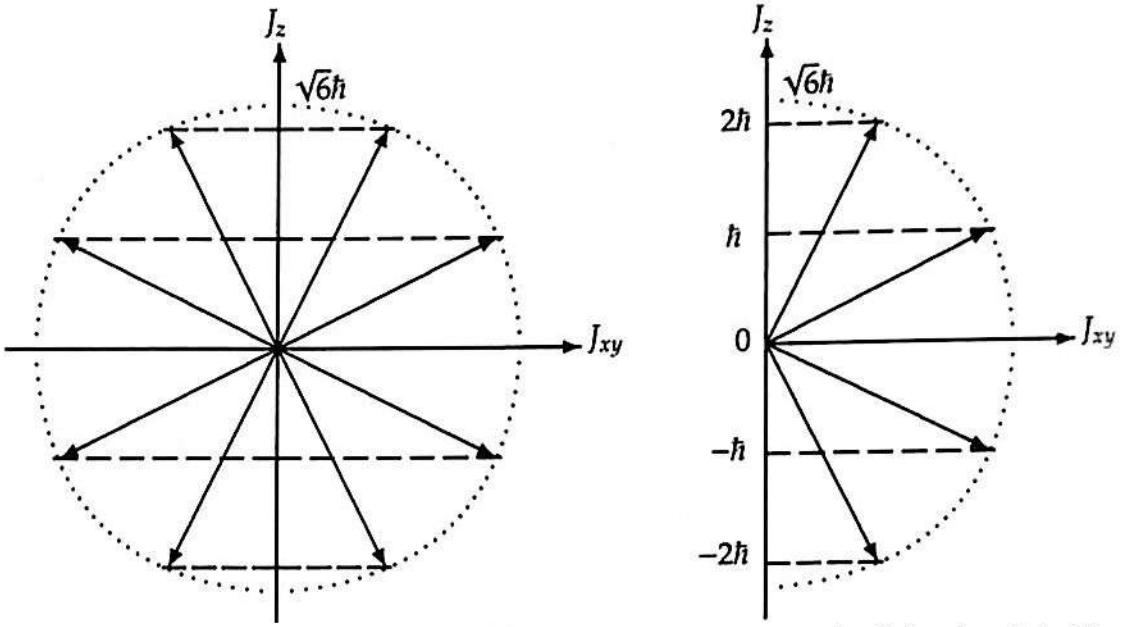


Figure 5.2: Graphical representation of the angular momentum $j = 2$ for the state $|\, 2, m\rangle$ with $m = -2, -1, 0, 1, 2$. The radius of the circle is $\hbar\sqrt{2(2+1)} = \sqrt{6}\hbar$.

5.6 Spin Angular Momentum

5.6.1 Experimental Evidence of the Spin

The history of the electron's spin is not quite linear; it is quite convoluted: while the first theoretical hypothesis about its existence was formulated in 1925 by Goudsmith and Uhlenbeck, the first experimental evidence of its existence was inadvertently found in 1922 by Stern and Gerlach, but they were not aware that their experiment provided a direct confirmation on the existence of the electron's spin, as we will see below. Actually, the story of the quantization of the electron's angular momentum began in 1896 with the Zeeman effect.

In studying the effect produced by a strong magnetic field on the light emitted by a radiant object, Pieter Zeeman[1] was the firs to have observed in 1896 the *splitting of a spectral line into several components* (of varying frequencies) as light enters a magnetic field. This effect later became known as the *Zeeman effect*.

Since the Zeeman effect takes place at the atomic scale and since classical physics (the only theoretical tool available to physicists around 1896) does not deal with atomic nor with subatomic phenomena, a satisfactory explanation of the Zeeman effect had to wait several more decades until the arrival of quantum mechanics. Around 1920, while the "old theory" of quantum mechanics was being developed by scientists such as Bohr and Sommerfeld, it was thought that the splitting of the atomic spectral lines is due to the *spatial* quantization of the electron's *orbital* angular momentum. As a result, Stern and Gerlach[2] designed and carried

[1] P. Zeeman, Reports of the Ordinary Sessions of the Mathematical and Physical Section (Royal Academy of Sciences in Amsterdam), 5, pp. 181–184 and 242–248 (1896); P. Zeeman, Philosophical Magazine 43 (262), pp. 226–239 (1897); P. Zeeman, Nature, 55 (1424), 347 (1897).

[2] W. Gerlach and O. Stern, Zeitschrift für Physik, 9 (1), pp. 349–352 (1922); W. Gerlach and O. Stern, Zeitschrift für Physik, 9 (1), pp. 353–355 (1922).

out an experiment in 1922 to test this very idea; namely, the Bohr-Sommerfeld hypothesis pertaining to the spatial quantization of the angular momentum of an atom's electron. In their experiment, Stern and Gerlach have found that, when a beam of silver atoms (Ag) passes through an inhomogeneous magnetic field, the beam *splits into two distinct components*. Although their experiment has showed clearly that spatial quantization exists (which offered experimental support to the Bohr-Sommerfeld quantization rule), they have not managed to explain why the beam splits into two components only. Later on, it was recognized that Stern and Gerlach had stumbled upon an important discovery: they have unwittingly provided the first experimental evidence that electrons have a *spin* angular momentum and that this spin is quantized and can take only two values ($\pm 1/2$ corresponding to the up and down values); hence, satisfactorily explaining why the silver beam splits into two components only. However, Stern and Gerlach were unaware of the importance of their discovery. As we will below, we had to wait three more years until two young graduate students, Goudsmith and Uhlenbeck, have introduced the concept of the *spin* in 1925.

So for about three years, scientists have tried in vain to understand the result of the Stern-Gerlach experiment — the splitting of a beam of silver atoms into two components. Silver has 47 electrons; 46 of them form a spherically symmetric charge distribution and the 47th electron occupies a 5s orbital. If the silver atom were in its ground state, its total orbital angular momentum would be zero: $l = 0$ (since the fifth shell electron would be in a 5s state). In the Stern–Gerlach experiment, a beam of silver atoms passes through an inhomogeneous (nonuniform) magnetic field. If, for argument's sake, the field were along the z-direction, we would expect classically to see on the screen a *continuous* band that is symmetric about the undeflected direction, $z = 0$. According to Schrödinger's wave theory, however, if the atoms had an orbital angular momentum l, we would expect the beam to split into an *odd* (*discrete*) number of $2l + 1$ components. Suppose the beam's atoms were in their ground state $l = 0$, there would be only one spot on the screen, and if the fifth shell electron were in a 5p state (i.e., $l = 1$), we would expect to see three spots. Experimentally, however, the beam behaves according to the predictions of neither classical physics nor Schrödinger's wave theory. Instead, *it splits into two distinct components* as shown in Figure 5.3a. This splitting into two distinct components was also observed for other hydrogen-like atoms in their ground states ($l = 0$), where no splitting is expected. So this puzzle remained an open question for several years.

To solve this puzzle as well as the Zeeman effect, Goudsmit and Uhlenbeck[3] hypothesized in 1925 that, in addition to its orbital angular momentum, the electron must possess an *intrinsic* angular momentum which, unlike the orbital angular momentum, has nothing to do with the spatial degrees of freedom. By analogy with the motion of the Earth, which consists of an orbital motion around the Sun and an internal rotational or *spinning* motion about its axis, the electron or, for that matter, any other microscopic particle may also be considered to have some sort of internal or intrinsic spinning motion. This intrinsic degree of freedom was given the suggestive name of *spin* angular momentum. One has to keep in mind, however, that the electron remains thus far a structureless or pointlike particle; hence caution has to be exercised when trying to link the electron's spin to an internal spinning motion. The spin angular momentum of a particle does not depend on its spatial degrees of freedom. The spin,

[3] S. Goudsmit and G.E. Uhlenbeck, Physica 5, 266 (1925); G.E. Uhlenbeck and S. Goudsmit, Naturwissenschaften 47, 953 (1925); G.E. Uhlenbeck and S. Goudsmit, Nature 117, 264 (1926); S. Goudsmit and G.E. Uhlenbeck, Physica 6, 273 (1926).

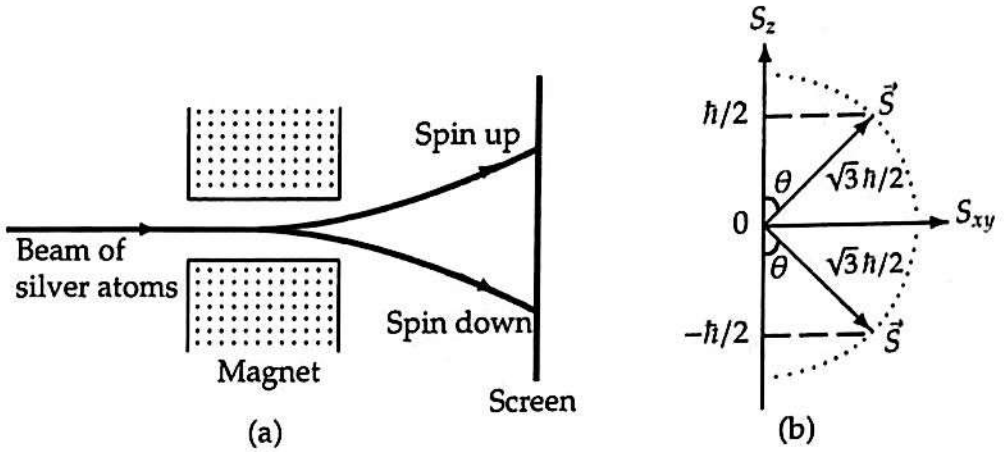


Figure 5.3: (a) Stern–Gerlach experiment: when a beam of silver atoms passes through an inhomogeneous magnetic field, it splits into two distinct components corresponding to spin-up and spin-down. (b) Graphical representation of spin $\frac{1}{2}$: the tip of $\vec{S}$ lies on a circle of radius $|\vec{S}| = \sqrt{3}\hbar/2$ so that its projection on the z-axis takes only two values, $\pm\hbar/2$, with $\theta = 54.73°$.

an intrinsic degree of freedom, is a purely quantum mechanical concept with no classical analog. Unlike the orbital angular momentum, *the spin cannot be described by a differential operator.*

From the classical theory of electromagnetism and as illustrated in Figure 5.4a, when a particle of rest mass m and charge q is rotating along a circle of radius r with a speed v, an electric current is generated: I = charge/(period) = q/T where the period for a full rotation is $T = (2\pi r)/v$ and, hence, the current is $I = qv/(2\pi r)$. The orbital angular momentum of this particle is $L = mvr$ and its *orbital magnetic dipole moment* μ is given by

$$\mu = (\text{Electric Current}) \cdot (\text{Loop Area}) = IA = \left(\frac{qv}{2\pi r}\right)(\pi r^2) = \frac{q}{2m}(mvr) \equiv \gamma L, \tag{5.90}$$

or by

$$\vec{\mu}_L = \frac{q}{2m}\vec{L} = \gamma_L \vec{L}, \qquad \gamma_L = \frac{q}{2m}, \tag{5.91}$$

where γ_L is the orbital *gyromagnetic ratio*; it is the ratio of the magnetic moment of the orbiting charged particle to its angular momentum: $\gamma_L = \mu/L$. As shown in Figure 5.4a, if the charge q is positive, $\vec{\mu}_L$ and $\vec{L}$ will be in the same direction; for a negative charge, such as an electron ($q = -e$), the magnetic dipole moment $\vec{\mu}_L = -e\vec{L}/(2m_e)$ and the orbital angular momentum will be in opposite directions. We should mention that the angular momentum $\vec{L}$ in (5.91) may be the orbital angular momentum, the spin angular momentum $\vec{S}$, or the total angular momentum $\vec{J}$. In fact, if we follow a classical analysis and picture the electron as a spinning spherical charge, and by analogy with (5.91), we can obtain an intrinsic or *spin magnetic dipole moment* for the electron:

$$\vec{\mu}_S = -\frac{e}{2m_e}\vec{S}. \tag{5.92}$$

This classical derivation of $\vec{\mu}_S$ is quite erroneous, since the electron cannot be viewed as a

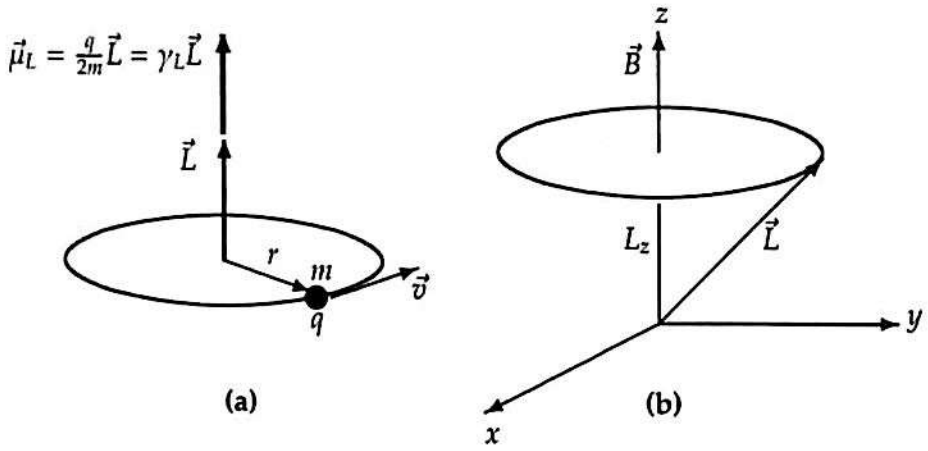


Figure 5.4: **(a):** Orbital magnetic dipole moment of a positive charge q orbiting with speed v along a circle of radius r is $\vec{\mu}_L = \gamma_L\vec{L}$ where $\gamma_L = q/(2m)$. **(b):** When an external magnetic field is applied, the orbital magnetic moment precesses about it.

spinning sphere; the electron is essentially considered to be point-like. Although the spin magnetic moment cannot be derived classically, as we did for the orbital magnetic moment, it can still be postulated by analogy with (5.91). However, it turns out that the classical result (5.92) is off by a factor g_s

$$\vec{\mu}_S = -g_s\frac{e}{2m_e}\vec{S} = \gamma_s\vec{S}, \qquad \gamma_s = -g_s\frac{e}{2m_e}, \tag{5.93}$$

where g_s is called the Landé factor or the spin *gyromagnetic factor*, or simply the g_s–factor, of the electron and γ_s is the spin gyromagnetic ratio; notice that the gyromagnetic factor is a dimensionless constant, while the gyromagnetic ratio has the dimensions of a charge over mass. We should note that both the Pauli and the Dirac equations give the correct value $g_s = 2$; in fact, we will show how to derive this value from the Pauli equation in Chapters 6 (see the Subsection about the Pauli equation) and from the Dirac equation in Chapter 12 (see Section about relativistic limit of the Dirac equation).

The spin magnetic moment is usually expressed in terms of the *Bohr magneton* μ_B as follows:

$$\vec{\mu}_S = \gamma_s\vec{S} \equiv -g_s\mu_B\frac{\vec{S}}{\hbar}, \qquad \gamma_s = -\frac{g_s\mu_B}{\hbar}, \qquad \mu_B = \frac{e\hbar}{2m_e} = 9.274\times10^{-24}\,\text{J/T}. \tag{5.94}$$

Clearly, the spin magnetic moment is quantized in units of μ_B, since the spin $\vec{S}$ is quantized in units of $\hbar$.

When the electron is placed in a magnetic field $\vec{B}$ and if the field is inhomogeneous, a force will be exerted on the electron's intrinsic dipole moment; the direction and magnitude of the force depend on the relative orientation of the field and the dipole. This force tends to align $\vec{\mu}_S$ along $\vec{B}$, producing a precessional motion of $\vec{\mu}_S$ around $\vec{B}$ (Figure 5.4b). For instance, if $\vec{\mu}_S$ is parallel to $\vec{B}$, the electron will move in the direction in which the field increases; conversely, if $\vec{\mu}_S$ is antiparallel to $\vec{B}$, the electron will move in the direction in which the field decreases. For hydrogen-like atoms (such as silver) that are in the ground state, the orbital angular momentum will be zero; hence the dipole moment of the atom will be entirely due to

the spin of the electron. The atomic beam will therefore deflect according to the orientation of the electron's spin. Since, experimentally, the beam splits into two components, the electron's spin must have only two possible orientations relative to the magnetic field, either parallel or antiparallel.

By analogy with the orbital angular momentum of a particle, which is characterized by two quantum numbers—the orbital number l and the azimuthal number m_l (with $m_l = -l, -l+1, \ldots, l-1, l$)—the spin angular momentum is also characterized by two quantum numbers, the spin s and its projection m_s on the z-axis (the direction of the magnetic field), where $m_s = -s, -s+1, \ldots, s-1, s$. Since only two components were observed in the Stern–Gerlach experiment, we must have $2s + 1 = 2$. The quantum numbers for the electron must then be given by $s = \frac{1}{2}$ and $m_s = \pm\frac{1}{2}$.

In nature it turns out that every fundamental particle has a specific spin. Some particles have integer spins $s = 0, 1, 2, \ldots$ (the pi mesons have spin $s = 0$, the photons have spin $s = 1$, and so on) and others have half-odd-integer spins $s = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$ (the electrons, protons, and neutrons have spin $s = \frac{1}{2}$, the deltas have spin $s = \frac{3}{2}$, and so on). We will see in Chapter 8 that particles with *half-odd-integer* spins are called *fermions* (quarks, electrons, protons, neutrons, etc.) and those with *integer* spins are called *bosons* (pions, photons, gravitons, etc.).

Besides confirming the existence of spin and measuring it, the Stern–Gerlach experiment offers a number of other important uses to quantum mechanics. First, by showing that a beam splits into a *discrete* set of components rather than a continuous band, it provides additional confirmation for the quantum hypothesis on the discrete character of the microphysical world. The Stern–Gerlach experiment also turns out to be an invaluable technique for preparing a quantum state. Suppose we want to prepare a beam of spin-up atoms; we simply pass an unpolarized beam through an inhomogeneous magnet, then collect the desired component and discard (or block) the other. The Stern–Gerlach experiment can also be used to determine the total angular momentum of an atom which, in the case where $l \neq 0$, is given by the sum of the orbital and spin angular momenta: $\vec{J} = \vec{L} + \vec{S}$. The addition of angular momenta is covered in Chapter 7.

In sum, while the first historical experimental evidence of the electron spin was due to the Stern-Gerlach experiment of 1922, the proper explanation of this experiment was given by Pauli in 1927 when he proposed his general theory of the spin.

5.6.2 General Theory of Spin

The theory of spin is identical to the general theory of angular momentum (Section 5.3). By analogy with the vector angular momentum $\hat{\vec{J}}$, the spin is also represented by a vector operator $\hat{\vec{S}}$ whose components $\hat{S}_x, \hat{S}_y, \hat{S}_z$ obey the same commutation relations as $\hat{J}_x, \hat{J}_y, \hat{J}_z$:

$$[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z, \qquad [\hat{S}_y, \hat{S}_z] = i\hbar\hat{S}_x, \qquad [\hat{S}_z, \hat{S}_x] = i\hbar\hat{S}_y. \tag{5.95}$$

In addition, $\hat{\vec{S}}^2$ and $\hat{S}_z$ commute; hence they have common eigenvectors:

$$\hat{\vec{S}}^2 \mid s, m_s\rangle = \hbar^2 s(s+1) \mid s, m_s\rangle, \qquad \hat{S}_z \mid s, m_s\rangle = \hbar m_s \mid s, m_s\rangle, \tag{5.96}$$

where $m_s = -s, -s+1, \ldots, s-1, s$. Similarly, we have

$$\hat{S}_\pm \mid s, m_s\rangle = \hbar\sqrt{s(s+1) - m_s(m_s \pm 1)} \mid s, m_s \pm 1\rangle, \tag{5.97}$$

where $\hat{S}_\pm = \hat{S}_x \pm i\hat{S}_y$, and

$$\langle \hat{S}_x^2 \rangle = \langle \hat{S}_y^2 \rangle = \frac{1}{2}(\langle \hat{\vec{S}}^2 \rangle - \langle \hat{S}_z^2 \rangle) = \frac{\hbar^2}{2}\left[s(s+1) - m_s^2\right], \tag{5.98}$$

where $\langle \hat{A} \rangle$ denotes $\langle \hat{A} \rangle = \langle s,\ m_s | \hat{A} \mid s,\ m_s \rangle$.

The spin states form an orthonormal and complete basis

$$\langle s',\ m_s' \mid s,\ m_s \rangle = \delta_{s',s}\delta_{m_s',m_s}, \qquad \sum_{m_s=-s}^{s} \mid s,\ m_s\rangle\langle s,\ m_s \mid = I, \tag{5.99}$$

where I is the unit matrix.

5.6.3 Spin 1/2 and the Pauli Matrices

For a particle with spin $\frac{1}{2}$ the quantum number m_s takes only two values: $m_s = -\frac{1}{2}$ and $\frac{1}{2}$. The particle can thus be found in either of the following two states: $\mid s,\ m_s\rangle = \left|\frac{1}{2},\ \frac{1}{2}\right\rangle$ and $\left|\frac{1}{2},\ -\frac{1}{2}\right\rangle$. The eigenvalues of $\hat{\vec{S}}^2$ and $\hat{S}_z$ are given by

$$\hat{\vec{S}}^2\left|\frac{1}{2},\ \pm\frac{1}{2}\right\rangle = \frac{3}{4}\hbar^2\left|\frac{1}{2},\ \pm\frac{1}{2}\right\rangle, \qquad \hat{S}_z\left|\frac{1}{2},\ \pm\frac{1}{2}\right\rangle = \pm\frac{\hbar}{2}\left|\frac{1}{2},\ \pm\frac{1}{2}\right\rangle. \tag{5.100}$$

Hence the spin may be represented graphically, as shown in Figure 5.3b, by a vector of length $|\vec{S}| = \sqrt{3}\hbar/2$, whose endpoint lies on a circle of radius $\sqrt{3}\hbar/2$, rotating along the surface of a cone with half-angle

$$\theta = \cos^{-1}\left(\frac{|m_s|}{\sqrt{s(s+1)}}\right) = \cos^{-1}\left(\frac{\hbar/2}{\sqrt{3}\hbar/2}\right) = \cos^{-1}\left(\frac{1}{\sqrt{3}}\right) = 54.73°. \tag{5.101}$$

The projection of $\hat{\vec{S}}$ on the z-axis is restricted to two values only: $\pm\hbar/2$ corresponding to spin-up and spin-down.

Let us now study the matrix representation of the spin $s = \frac{1}{2}$. Using (5.67) and (5.68) we can represent the operators $\hat{\vec{S}}^2$ and $\hat{S}_z$ within the $\{|s,\ m_s\rangle\}$ basis by the following matrices:

$$\hat{\vec{S}}^2 = \begin{pmatrix} \langle \frac{1}{2},\ \frac{1}{2} \mid \hat{\vec{S}}^2 \mid \frac{1}{2},\ \frac{1}{2}\rangle & \langle \frac{1}{2},\ \frac{1}{2} \mid \hat{\vec{S}}^2 \mid \frac{1}{2},\ -\frac{1}{2}\rangle \\ \langle \frac{1}{2},\ -\frac{1}{2} \mid \hat{\vec{S}}^2 \mid \frac{1}{2},\ \frac{1}{2}\rangle & \langle \frac{1}{2},\ -\frac{1}{2} \mid \hat{\vec{S}}^2 \mid \frac{1}{2},\ -\frac{1}{2}\rangle \end{pmatrix} = \frac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{5.102}$$

$$\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.103}$$

The matrices of $\hat{S}_+$ and $\hat{S}_-$ can be inferred from (5.69):

$$\hat{S}_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \hat{S}_- = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{5.104}$$

and since $\hat{S}_x = \frac{1}{2}(\hat{S}_+ + \hat{S}_-)$ and $\hat{S}_y = \frac{i}{2}(\hat{S}_- - \hat{S}_+)$, we have

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{5.105}$$

The joint eigenvectors of $\hat{\vec{S}}^2$ and $\hat{S}_z$ are expressed in terms of two-element column matrices, known as the *Pauli spinors*, that have eigenvalues $S_z = \pm\hbar/2$:

$$\boxed{\left|\tfrac{1}{2}, \tfrac{1}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ eigenvalue } S_z = +\tfrac{\hbar}{2}, \qquad \left|\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{ eigenvalue } S_z = -\tfrac{\hbar}{2}.} \tag{5.106}$$

It is easy to verify that these eigenvectors form a basis that is complete,

$$\sum_{m_s=-\frac{1}{2}}^{\frac{1}{2}} \left|\frac{1}{2}, m_s\right\rangle\left\langle\frac{1}{2}, m_s\right| = \begin{pmatrix} 0 \\ 1 \end{pmatrix}(0\ 1) + \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1\ 0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{5.107}$$

and orthonormal,

$$\left\langle\frac{1}{2}, \frac{1}{2}\middle|\frac{1}{2}, \frac{1}{2}\right\rangle = (1\ 0)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1, \tag{5.108}$$

$$\left\langle\frac{1}{2}, -\frac{1}{2}\middle|\frac{1}{2}, -\frac{1}{2}\right\rangle = (0\ 1)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1, \tag{5.109}$$

$$\left\langle\frac{1}{2}, \frac{1}{2}\middle|\frac{1}{2}, -\frac{1}{2}\right\rangle = \left\langle\frac{1}{2}, -\frac{1}{2}\middle|\frac{1}{2}, \frac{1}{2}\right\rangle = 0. \tag{5.110}$$

Let us now find the eigenvectors of $\hat{S}_x$ and $\hat{S}_y$. First, note that the basis vectors $|s, m_s\rangle$ are eigenvectors of neither $\hat{S}_x$ nor $\hat{S}_y$; their eigenvectors can, however, be expressed in terms of $|s, m_s\rangle$ as follows:

$$|\psi_x\rangle_\pm = \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2}, \frac{1}{2}\right\rangle \pm \left|\frac{1}{2}, -\frac{1}{2}\right\rangle\right], \tag{5.111}$$

$$|\psi_y\rangle_\pm = \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2}, \frac{1}{2}\right\rangle \pm i\left|\frac{1}{2}, -\frac{1}{2}\right\rangle\right]. \tag{5.112}$$

The eigenvalue equations for $\hat{S}_x$ and $\hat{S}_y$ are thus given by

$$\hat{S}_x|\psi_x\rangle_\pm = \pm\frac{\hbar}{2}|\psi_x\rangle_\pm, \qquad \hat{S}_y|\psi_y\rangle_\pm = \pm\frac{\hbar}{2}|\psi_y\rangle_\pm. \tag{5.113}$$

Pauli matrices

When $s = \frac{1}{2}$ it is convenient to introduce the *Pauli matrices* $\sigma_x, \sigma_y, \sigma_z$, which are related to the spin vector as follows:

$$\hat{\vec{S}} = \frac{\hbar}{2}\vec{\sigma}. \tag{5.114}$$

Using this relation along with (5.103) and (5.105), we have

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.115}$$

These matrices satisfy the following two properties:

$$\sigma_j^2 = \hat{I} \qquad (j = x, y, z), \tag{5.116}$$

$$\sigma_j \sigma_k + \sigma_k \sigma_j = 0 \qquad (j \neq k), \tag{5.117}$$

where the subscripts j and k refer to x, y, z, and $\hat{I}$ is the 2×2 unit matrix. These two equations are equivalent to the anticommutation relation

$$\{\sigma_j, \sigma_k\} = 2\hat{I}\delta_{j,k}, \qquad (j, k = x, y, z). \tag{5.118}$$

This anticommutation relation defines what is known as the three-dimensional Clifford algebra

We can verify that the Pauli matrices satisfy the commutation relations

$$[\sigma_j, \sigma_k] = 2i\, \varepsilon_{jkl} \sigma_l, \tag{5.119}$$

where ε_{jkl} is the antisymmetric tensor (also known as the Levi–Civita tensor)

$$\varepsilon_{jkl} = \begin{cases} 1 & \text{for } j, k, l \text{ an even permutation of } x, y, z, \\ 0 & \text{if any two or more indices among } j, k, l \text{ are equal,} \\ -1 & \text{for } j, k, l \text{ an odd permutation of } x, y, z. \end{cases} \tag{5.120}$$

We can condense the relations (5.116), (5.117), and (5.119) into

$$\sigma_j \sigma_k = \delta_{j,k} + i \sum_l \varepsilon_{jkl} \sigma_l. \tag{5.121}$$

Using this relation we can verify that, for any two arbitrary vectors $\vec{A}$ and $\vec{B}$ which commute with $\vec{\sigma}$, we have

$$(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) = (\vec{A} \cdot \vec{B})\hat{I} + i\vec{\sigma} \cdot (\vec{A} \times \vec{B}), \tag{5.122}$$

where $\hat{I}$ is the 2×2 unit matrix. The Pauli matrices are Hermitian, traceless, and have determinants equal to -1:

$$\sigma_j^\dagger = \sigma_j, \qquad \mathrm{Tr}(\sigma_j) = 0, \qquad \det(\sigma_j) = -1 \qquad (j = x, y, z). \tag{5.123}$$

Using the relation $\sigma_x \sigma_y = i\sigma_z$ along with $\sigma_z^2 = \hat{I}$, we obtain

$$\sigma_x \sigma_y \sigma_z = i\hat{I}. \tag{5.124}$$

From the commutation relations (5.119) we can show that

$$e^{i\alpha\sigma_j} = I \cos\alpha + i\sigma_j \sin\alpha \qquad (j = x, y, z), \tag{5.125}$$

where I is the unit matrix and α is an arbitrary real constant.

Example 5.4
Find the energy levels of a spin $s = \frac{3}{2}$ particle whose Hamiltonian is given by

$$\hat{H} = \frac{\alpha}{\hbar^2}(\hat{S}_x^2 + \hat{S}_y^2 - 2\hat{S}_z^2) - \frac{\beta}{\hbar}\hat{S}_z;$$

α and β are constants. Are these levels degenerate?

Solution
Rewriting $\hat{H}$ in the form,

$$\hat{H} = \frac{\alpha}{\hbar^2}\left(\hat{\vec{S}}^2 - 3\hat{S}_z^2\right) - \frac{\beta}{\hbar}\hat{S}_z, \tag{5.126}$$

we see that $\hat{H}$ is diagonal in the $\{|s, m\rangle\}$ basis:

$$E_m = \langle s, m \mid \hat{H} \mid s, m\rangle = \frac{\alpha}{\hbar^2}\left[\hbar^2 s(s+1) - 3\hbar^2 m^2\right] - \frac{\beta}{\hbar}\hbar m = \frac{15}{4}\alpha - m(3\alpha m + \beta), \tag{5.127}$$

where the quantum number m takes any of the four values $m = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$. Since E_m depends on m, the energy levels of this particle are nondegenerate.

5.6.4 Schrödinger Equation with Spin: Pauli Hamiltonian

Up to now, we have dealt with spinless quantum mechanics only. In its original form, the Schrödinger equation does not deal with spin, it deals only with the motion of quantum particles in space. Since the spin does not depend on the spatial degrees of freedom, the components $\hat{S}_x, \hat{S}_y, \hat{S}_z$ of the spin operator commute with all the spatial operators, notably the orbital angular momentum $\hat{\vec{L}}$, the position and the momentum operators $\hat{\vec{R}}$ and $\hat{\vec{P}}$:

$$[\hat{S}_j, \hat{L}_k] = 0, \qquad [\hat{S}_j, \hat{R}_k] = 0, \qquad [\hat{S}_j, \hat{P}_k] = 0 \qquad (j,k = x,y,z). \tag{5.128}$$

Clearly, without the inclusion of the internal degrees of freedom of particles, spinless quantum mechanics, which is founded on the Schrödinger equation, is incomplete. A complete description requires the addition of spin. As evidenced by the Stern-Gerlach experiment, in the absence of an external magnetic field, the two states of a pin 1/2 system—spin-up and spin-down states—are undistinguishable and have the same energy (degenerate). So, in the absence of an external magnetic field, the system can be described correctly by the spinless Schrödinger equation. However, in the presence of a magnetic field, these two states split and become distinguishable, but the Schrödinger equation provides no mechanism to account for this splitting unless one has to find a way to add the spin degrees of freedom and that the state of the system must be a linear superposition of all the components of these internal degrees of freedom. In this case, the total wave function $|\Psi\rangle$ must consist of a product of two parts: a spatial part $\psi(\vec{r})$ and a spin part $|s, m_s\rangle$:

$$|\Psi\rangle = |\psi\rangle\,|s, m_s\rangle. \tag{5.129}$$

This product of the space and spin degrees of freedom is not a product in the usual sense, because they belong to different spaces; the product between degrees of freedom belonging

to different spaces can be handled properly by means of a tensor product; we will discuss this issue in sufficient detail in Chapter 7. Additionally, we will show in Chapter 6 that the four quantum numbers n, l, m_l, and m_s are required to completely describe the state of an electron moving in a central field; its wave function is

$$\Psi_{nlm_lm_s}(\vec{r}) = \psi_{nlm_l}(\vec{r}) \mid s,\ m_s\rangle. \tag{5.130}$$

Since the spin operator does not depend on the spatial degrees of freedom, it acts only on the spin part $\mid s,\ m_s\rangle$ and leaving the spatial wave function, $\psi_{nlm_l}(\vec{r})$, unchanged; conversely, the spatial operators $\hat{\vec{L}}$, $\hat{\vec{R}}$, and $\hat{\vec{P}}$ act on the spatial part and not on the spin part.

For spin 1/2 particles, using the spin-up and spin-down spinors (5.106),

$$\Psi_{nlm_l\frac{1}{2}}(\vec{r}) = \psi_1(\vec{r})\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\psi_1(\vec{r})\\0\end{pmatrix}, \qquad \Psi_{nlm_l-\frac{1}{2}}(\vec{r}) = \psi_2(\vec{r})\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\\psi_2(\vec{r})\end{pmatrix}, \tag{5.131}$$

we can write the total wave function $\Psi(\vec{r})$ as a linear superposition as follows

$$\Psi(\vec{r}) = \Psi_{nlm_l\frac{1}{2}}(\vec{r}) + \Psi_{nlm_l-\frac{1}{2}}(\vec{r}) = \begin{pmatrix}\psi_1(\vec{r})\\0\end{pmatrix} + \begin{pmatrix}0\\\psi_2(\vec{r})\end{pmatrix} = \begin{pmatrix}\psi_1(\vec{r})\\\psi_2(\vec{r})\end{pmatrix}, \tag{5.132}$$

or, in terms of the Dirac bracket notation, as

$$\mid \Psi\rangle = \mid \psi_1\rangle\left|\frac{1}{2},\ \frac{1}{2}\right\rangle + \mid \psi_2\rangle\left|\frac{1}{2},\ -\frac{1}{2}\right\rangle = \mid \psi_1\rangle\begin{pmatrix}1\\0\end{pmatrix} + \mid \psi_2\rangle\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}\mid \psi_1\rangle\\\mid \psi_2\rangle\end{pmatrix}. \tag{5.133}$$

It is important to contrast the Schrödinger and Pauli descriptions of quantum systems. On the one hand, in the Schrödinger description, a system is described by means of the standard *scalar* Schrödinger equation where the wave function is a *single component* scalar function (wave function), $\psi(\vec{r})$, and the operators (such as the Hamiltonian) are also *single component* differential operators: $\hat{H}_S = \hat{\vec{P}}^2/(2m) + \hat{V}(\vec{r})$. In contrast, in the Pauli description, the wave function of a spin 1/2 particle is described by means of a 2×1 column matrix (the Pauli spinor) with complex components, while the operators must be represented by 2×2 matrices. That is, instead of the single component Schrödinger Hamiltonian, $\hat{H}_S$, we will have a 2×2 diagonal Hamiltonian:

$$\hat{H}_P = \begin{pmatrix}\hat{\vec{P}}^2/(2m) + \hat{V}(\vec{r}) & 0\\ 0 & \hat{\vec{P}}^2/(2m) + \hat{V}(\vec{r})\end{pmatrix} = \left(\frac{\hat{\vec{P}}^2}{2m} + \hat{V}\right)\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} = \hat{H}_S\hat{I}_2, \tag{5.134}$$

where $\hat{I}_2$ is the 2×2 unit matrix; $\hat{H}_P$ is known as the *Pauli Hamiltonian* for a spin 1/2 particle. Additionally, we will derive the Pauli Hamiltonian for a spin 1/2 particle in an external magnetic field in Chapter 6 (see equation (12.475) on Page 796). So, in the absence of an external magnetic field, the Pauli equation

$$\hat{H}_P \mid \Psi_P\rangle = \begin{pmatrix}\hat{\vec{P}}^2/(2m) + \hat{V}(\vec{r}) & 0\\ 0 & \hat{\vec{P}}^2/(2m) + \hat{V}(\vec{r})\end{pmatrix} \mid \Psi_P\rangle, \tag{5.135}$$

will consist of two identical decoupled Schrödinger equations in the standard form:

$$\hat{H}_S \mid \psi\rangle = \left(\frac{\hat{\vec{P}}^2}{2m} + \hat{V}(\vec{r})\right) \mid \psi\rangle, \tag{5.136}$$

where the spinor $|\Psi_P\rangle$ is a 2×1 column matrix, while $|\psi\rangle$ is a single component scalar state.

So, in the Pauli formalism, the wave function of a spin 1/2 particle is represented by a 2×1 column matrix and operators by 2×2 matrices, whereas the Schrödinger formalism deals only with "single component" quantities—wave functions and operators. We should note that, when we deal with relativistic quantum mechanics in Chapter 12, we will see that the Dirac equation describes electrons by means of 4×1 column spinors and the operators by 4×4 matrices. Self-evidently, since the Dirac spinor jointly describes a spin 1/2 particle and its antiparticle, we would need *4 components* because two components are needed to account for the spin-up and spin-down states of the *particle*, while the remaining two components are needed for the spin-up and spin-down states of the *antiparticle*. However, since the Pauli formalism deals with particles only (no antiparticles), two component spinors and 2×2 operators will be sufficient to account for the spinors and observables relevant to spin 1/2 particles.

5.7 Eigenfunctions of the Orbital Angular Momentum

We now turn to the coordinate representation of the angular momentum. In this section, we are going to work within the spherical coordinate system. Let us denote the joint eigenstates of $\hat{\vec{L}}^2$ and $\hat{L}_z$ by $|l, m\rangle$:

$$\boxed{\hat{\vec{L}}^2 |l, m\rangle = \hbar^2 l(l+1) |l, m\rangle,} \tag{5.137}$$

$$\boxed{\hat{L}_z |l, m\rangle = \hbar m |l, m\rangle.} \tag{5.138}$$

The operators $\hat{L}_z$, $\hat{\vec{L}}^2$, $\hat{L}_\pm$, whose Cartesian components are listed in Eqs (5.3) to (5.5), can be expressed in terms of spherical coordinates (Appendix B) as follows:

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \varphi}, \tag{5.139}$$

$$\hat{\vec{L}}^2 = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \varphi^2} \right], \tag{5.140}$$

$$\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y = \pm\hbar e^{\pm i\varphi} \left[\frac{\partial}{\partial \theta} \pm i \frac{\cos\theta}{\sin\theta} \frac{\partial}{\partial \varphi} \right]. \tag{5.141}$$

Since the operators $\hat{L}_z$ and $\hat{\vec{L}}$ depend only on the angles θ and φ, their eigenstates depend only on θ and φ. Denoting their joint eigenstates by

$$\boxed{\langle \theta\varphi |\, l, m\rangle = Y_{lm}(\theta, \varphi),} \tag{5.142}$$

where[4] $Y_{lm}(\theta, \varphi)$ are continuous functions of θ and φ, we can rewrite the eigenvalue equations (5.137) and (5.138) as follows:

$$\boxed{\hat{\vec{L}}^2 Y_{lm}(\theta, \varphi) = \hbar^2 l(l+1) Y_{lm}(\theta, \varphi),} \tag{5.143}$$

[4]For notational consistency throughout this text, we will insert a comma between l and m in $Y_{lm}(\theta, \varphi)$ whenever m is negative.

$$\boxed{\hat{L}_z Y_{lm}(\theta,\varphi) = m\hbar Y_{lm}(\theta,\varphi).} \tag{5.144}$$

Since $\hat{L}_z$ depends only on φ, as shown in (5.139), the previous two equations suggest that the eigenfunctions $Y_{lm}(\theta,\varphi)$ are separable:

$$Y_{lm}(\theta,\varphi) = \Theta_{lm}(\theta)\Phi_m(\varphi). \tag{5.145}$$

We ascertain that

$$\boxed{\hat{L}_\pm Y_{lm}(\theta,\varphi) = \hbar\sqrt{l(l+1) - m(m\pm 1)}\, Y_{l\,m\pm 1}(\theta,\varphi).} \tag{5.146}$$

5.7.1 Eigenfunctions and Eigenvalues of $\hat{L}_z$

Inserting (5.145) into (5.144) we obtain $\hat{L}_z\Theta_{lm}(\theta)\Phi_m(\varphi) = m\hbar\Theta_{lm}(\theta)\Phi_m(\varphi)$. Now since $\hat{L}_z = -i\hbar\partial/\partial\varphi$, we have

$$-i\hbar\Theta_{lm}(\theta)\frac{\partial\Phi_m(\varphi)}{\partial\varphi} = m\hbar\Theta_{lm}(\theta)\Phi_m(\varphi), \tag{5.147}$$

which reduces to

$$-i\frac{\partial\Phi_m(\varphi)}{\partial\varphi} = m\Phi_m(\varphi). \tag{5.148}$$

The normalized solutions of this equation are given by

$$\Phi_m(\varphi) = \frac{1}{\sqrt{2\pi}}e^{im\varphi}, \tag{5.149}$$

where $1/\sqrt{2\pi}$ is the normalization constant,

$$\int_0^{2\pi} d\varphi\, \Phi^*_{m'}(\varphi)\Phi_m(\varphi) = \delta_{m',m}. \tag{5.150}$$

For $\Phi_m(\varphi)$ to be single-valued, it must be periodic in φ with period 2π, $\Phi_m(\varphi + 2\pi) = \Phi_m(\varphi)$; hence

$$e^{im(\varphi+2\pi)} = e^{im\varphi}. \tag{5.151}$$

This relation shows that the expectation value of $\hat{L}_z$, $l_z = \langle l,\ m \mid \hat{L}_z \mid l\ m\rangle$, is restricted to a *discrete* set of values

$$l_z = m\hbar, \qquad m = 0, \pm1, \pm2, \pm3, \ldots. \tag{5.152}$$

Thus, the values of m vary from $-l$ to l:

$$m = -l, -(l-1), -(l-2), \ldots, 0, 1, 2, \ldots, l-2, l-1, l. \tag{5.153}$$

Hence the quantum number l must also be an integer. This is expected since the orbital angular momentum must have integer values.

5.7.2 Eigenfunctions of $\hat{\vec{L}}^2$

Let us now focus on determining the eigenfunctions $\Theta_{lm}(\theta)$ of $\hat{\vec{L}}^2$. We are going to follow two methods. The first method involves differential equations and gives $\Theta_{lm}(\theta)$ in terms of the well-known associated Legendre functions. The second method is algebraic; it deals with the operators $\hat{L}_\pm$ and enables an explicit construction of $Y_{lm}(\theta,\varphi)$, the spherical harmonics.

5.7.2.1 First Method for Determining the Eigenfunctions of $\hat{L}^2$

We begin by applying $\hat{L}^2$ of (5.140) to the eigenfunctions

$$Y_{lm}(\theta,\varphi) = \frac{1}{\sqrt{2\pi}}\Theta_{lm}(\theta)e^{im\varphi}. \tag{5.154}$$

This gives

$$\begin{aligned}\hat{L}^2 Y_{lm}(\theta,\varphi) &= \frac{-\hbar^2}{\sqrt{2\pi}}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]\Theta_{lm}(\theta)e^{im\varphi}\\ &= \frac{\hbar^2 l(l+1)}{\sqrt{2\pi}}\Theta_{lm}(\theta)e^{im\varphi},\end{aligned} \tag{5.155}$$

which, after eliminating the φ-dependence, reduces to

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta_{lm}(\theta)}{d\theta}\right)+\left[l(l+1)-\frac{m^2}{\sin^2\theta}\right]\Theta_{lm}(\theta)=0. \tag{5.156}$$

This equation is known as the *Legendre differential equation*. Its solutions can be expressed in terms of the *associated Legendre functions* $P_l^m(\cos\theta)$:

$$\Theta_{lm}(\theta) = C_{lm}P_l^m(\cos\theta), \tag{5.157}$$

which are defined by

$$P_l^m(x) = (1-x^2)^{|m|/2}\frac{d^{|m|}}{dx^{|m|}}P_l(x). \tag{5.158}$$

This shows that

$$P_l^{-m}(x) = P_l^m(x), \tag{5.159}$$

where $P_l(x)$ is the lth *Legendre polynomial* which is defined by the *Rodrigues formula*

$$P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2-1)^l. \tag{5.160}$$

We can obtain at once the first few Legendre polynomials:

$$P_0(x) = 1, \qquad P_1(x) = \frac{1}{2}\frac{d(x^2-1)}{dx} = x, \tag{5.161}$$

$$P_2(x) = \frac{1}{8}\frac{d^2(x^2-1)^2}{dx^2} = \frac{1}{2}(3x^2-1), \qquad P_3(x) = \frac{1}{48}\frac{d^3(x^2-1)^3}{dx^3} = \frac{1}{2}(5x^3-3x), \tag{5.162}$$

$$P_4(x) = \frac{1}{8}(35x^4-30x^2+3), \qquad P_5(x) = \frac{1}{8}(63x^5-70x^3+15x). \tag{5.163}$$

The Legendre polynomials satisfy the following closure or completeness relation:

$$\frac{1}{2}\sum_{l=0}^{\infty}(2l+1)P_l(x')P_l(x) = \delta(x-x'). \tag{5.164}$$

Table 5.1: First few Legendre polynomials and associated Legendre functions.

Legendre polynomials	Associated Legendre functions
$P_0(\cos\theta) = 1$	$P_1^1(\cos\theta) = \sin\theta$
$P_1(\cos\theta) = \cos\theta$	$P_2^1(\cos\theta) = 3\cos\theta\sin\theta$
$P_2(\cos\theta) = \frac{1}{2}(3\cos^2\theta - 1)$	$P_2^2(\cos\theta) = 3\sin^2\theta$
$P_3(\cos\theta) = \frac{1}{2}(5\cos^3\theta - 3\cos\theta)$	$P_3^1(\cos\theta) = \frac{3}{2}\sin\theta(5\cos^2\theta - 1)$
$P_4(\cos\theta) = \frac{1}{8}(35\cos^4\theta - 30\cos^2\theta + 3)$	$P_3^2(\cos\theta) = 15\sin^2\theta\cos\theta$
$P_5(\cos\theta) = \frac{1}{8}(63\cos^5\theta - 70\cos^3\theta + 15\cos\theta)$	$P_3^3(\cos\theta) = 15\sin^3\theta$

From (5.160) we can infer at once

$$P_l(-x) = (-1)^l P_l(x). \tag{5.165}$$

A similar calculation leads to the first few associated Legendre functions:

$$P_1^1(x) = \sqrt{1-x^2}, \tag{5.166}$$

$$P_2^1(x) = 3x\sqrt{1-x^2}, \qquad P_2^2(x) = 3(1-x^2), \tag{5.167}$$

$$P_3^1(x) = \frac{3}{2}(5x^2-1)\sqrt{1-x^2}, \quad P_3^2(x) = 15x(1-x^2), \quad P_3^3(x) = 15(1-x^2)^{3/2}, \tag{5.168}$$

where $P_l^0(x) = P_l(x)$, with $l = 0, 1, 2, 3, \ldots$. The first few expressions for the associated Legendre functions and the Legendre polynomials are listed in Table 5.1. Note that

$$P_l^m(-x) = (-1)^{l+m} P_l^m(x). \tag{5.169}$$

The constant C_{lm} of (5.157) can be determined from the orthonormalization condition

$$\langle l',\, m' \mid l,\, m\rangle = \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta\, \sin\theta \langle l', m'|\theta\,\varphi\rangle\langle\theta\varphi \mid l,\, m\rangle = \delta_{l',l}\delta_{m',m}, \tag{5.170}$$

which can be written as

$$\boxed{\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta\, \sin\theta\, Y^*_{l'm'}(\theta,\varphi) Y_{lm}(\theta,\varphi) = \delta_{l',l}\delta_{m',m}.} \tag{5.171}$$

This relation is known as the normalization condition of spherical harmonics. Using the form (5.154) for $Y_{lm}(\theta,\varphi)$, we obtain

$$\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta\, \sin\theta \left|Y_{lm}(\theta,\varphi)\right|^2 = \frac{|C_{lm}|^2}{2\pi} \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta |P_l^m(\cos\theta)|^2 = 1. \tag{5.172}$$

From the theory of associated Legendre functions, we have

$$\int_0^{\pi} d\theta\, \sin\theta P_l^m(\cos\theta) P_{l'}^m(\cos\theta) = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\delta_{l,l'}, \tag{5.173}$$

which is known as the normalization condition of associated Legendre functions. A combination of the previous two relations leads to an expression for the coefficient C_{lm}:

$$C_{lm} = (-1)^m \sqrt{\left(\frac{2l+1}{2}\right)\frac{(l-m)!}{(l+m)!}} \qquad (m \geq 0). \tag{5.174}$$

Inserting this equation into (5.157), we obtain the eigenfunctions of $\hat{\vec{L}}^2$:

$$\Theta_{lm}(\theta) = (-1)^m \sqrt{\left(\frac{2l+1}{2}\right)\frac{(l-m)!}{(l+m)!}}\, P_l^m(\cos\theta). \tag{5.175}$$

Finally, the joint eigenfunctions, $Y_{lm}(\theta,\varphi)$, of $\hat{\vec{L}}^2$ and $\hat{L}_z$ can be obtained by substituting (5.149) and (5.175) into (5.145):

$$\boxed{Y_{lm}(\theta,\varphi) = (-1)^m \sqrt{\left(\frac{2l+1}{4\pi}\right)\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta)e^{im\varphi} \qquad (m \geq 0).} \tag{5.176}$$

These are called the *normalized spherical harmonics*.

5.7.2.2 Second Method for Determining the Eigenfunctions of $\hat{\vec{L}}^2$

The second method deals with a direct construction of $Y_{lm}(\theta,\varphi)$; it starts with the case $m = l$ (this is the maximum value of m). By analogy with the general angular momentum algebra developed in the previous section, the action of $\hat{L}_+$ on Y_{ll} gives zero,

$$\langle\theta\,\varphi|\hat{L}_+ \mid l,\, l\rangle = \hat{L}_+ Y_{ll}(\theta,\varphi) = 0, \tag{5.177}$$

since Y_{ll} cannot be raised further as $Y_{ll} = Y_{lm_{max}}$.

Using the expression (5.141) for $\hat{L}_+$ in the spherical coordinates, we can rewrite (5.177) as follows:

$$\frac{\hbar e^{i\varphi}}{\sqrt{2\pi}}\left[\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right]\Theta_{ll}(\theta)e^{il\varphi} = 0, \tag{5.178}$$

which leads to

$$\frac{1}{\Theta_{ll}}\frac{\partial\Theta_{ll}(\theta)}{\partial\theta} = l\cot\theta. \tag{5.179}$$

The solution to this differential equation is of the form

$$\Theta_{ll}(\theta) = C_l \sin^l\theta, \tag{5.180}$$

where C_l is a constant to be determined from the normalization condition (5.171) of $Y_{ll}(\theta,\varphi)$:

$$Y_{ll}(\theta,\varphi) = \frac{C_l}{\sqrt{2\pi}}e^{il\varphi}\sin^l\theta. \tag{5.181}$$

We can ascertain that C_l is given by

$$C_l = \frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l+1)!}{2}}. \tag{5.182}$$

The action of $\hat{L}_-$ on $Y_{ll}(\theta,\varphi)$ is given, on the one hand, by

$$\hat{L}_- Y_{ll}(\theta,\varphi) = \hbar\sqrt{2l}\,Y_{l,l-1}(\theta,\varphi) \tag{5.183}$$

and, on the other hand, by

$$\hat{L}_- Y_{ll}(\theta,\varphi) = \hbar\frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}}e^{i(l-1)\varphi}(\sin\theta)^{1-l}\frac{d}{d(\cos\theta)}[(\sin\theta)^{2l}], \tag{5.184}$$

where we have used the spherical coordinate form (5.141).

Similarly, we can show that the action of $\hat{L}_-^{l-m}$ on $Y_{ll}(\theta,\varphi)$ is given, on the one hand, by

$$\hat{L}_-^{l-m} Y_{ll}(\theta,\varphi) = \hbar^{l-m}\sqrt{\frac{(2l)!(l+m)!}{(l-m)!}}Y_{lm}(\theta,\varphi) \tag{5.185}$$

and, on the other hand, by

$$\hat{L}_-^{l-m} Y_{ll}(\theta,\varphi) = \hbar^{l-m}\frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l)!(2l+1)!}{4\pi}}e^{im\varphi}\frac{1}{\sin^m\theta}\frac{d^{l-m}}{d(\cos\theta)^{l-m}}(\sin\theta)^{2l}, \tag{5.186}$$

where $m \geq 0$. Equating the previous two relations, we obtain the expression of the spherical harmonic $Y_{lm}(\theta,\varphi)$ for $m \geq 0$:

$$\boxed{Y_{lm}(\theta,\varphi) = \frac{(-1)^l}{2^l l!}\sqrt{\left(\frac{2l+1}{4\pi}\right)\frac{(l+m)!}{(l-m)!}}e^{im\varphi}\frac{1}{\sin^m\theta}\frac{d^{l-m}}{d(\cos\theta)^{l-m}}(\sin\theta)^{2l}.} \tag{5.187}$$

5.7.3 Properties of the Spherical Harmonics

Since the spherical harmonics $Y_{lm}(\theta,\varphi)$ are joint eigenfunctions of $\hat{\vec{L}}^2$ and $\hat{L}_z$ and are orthonormal (5.171), they constitute an orthonormal basis in the Hilbert space of square-integrable functions of θ and φ. The completeness relation is given by

$$\sum_{m=-l}^{l} |\, l,\, m\rangle\langle l,\, m\,| = 1 \tag{5.188}$$

or

$$\begin{aligned}\sum_m \langle\theta\varphi\,|\,l,\,m\rangle\langle l,\,m\,|\,\theta'\varphi'\rangle &= \sum_m Y_{lm}^*(\theta',\varphi')Y_{lm}(\theta,\varphi) = \delta(\cos\theta-\cos\theta')\delta(\varphi-\varphi') \\ &= \frac{\delta(\theta-\theta')}{\sin\theta}\delta(\varphi-\varphi').\end{aligned} \tag{5.189}$$

Let us mention some essential properties of the spherical harmonics. First, the spherical harmonics are complex functions; their complex conjugate is given by

$$\boxed{[Y_{lm}(\theta,\varphi)]^* = (-1)^m Y_{l,-m}(\theta,\varphi).} \tag{5.190}$$

We can verify that $Y_{lm}(\theta,\varphi)$ is an eigenstate of the parity operator $\hat{\mathcal{P}}$ with an eigenvalue $(-1)^l$:

$$\hat{\mathcal{P}}Y_{lm}(\theta,\varphi) = Y_{lm}(\pi-\theta,\varphi+\pi) = (-1)^l Y_{lm}(\theta,\varphi), \tag{5.191}$$

since a spatial reflection about the origin, $\vec{r}' = -\vec{r}$, corresponds to $r' = r$, $\theta' = \pi - \theta$, and $\varphi' = \pi + \varphi$, which leads to $P_l^m(\cos\theta') = P_l^m(-\cos\theta) = (-1)^{l+m}P_l^m(\cos\theta)$ and $e^{im\varphi'} = e^{im\pi}e^{im\varphi} = (-1)^m e^{im\varphi}$. In this context, we may recall that the parity operation causes a reflection of all the coordinates through the origin $\vec{r} \longrightarrow \vec{r}' = -\vec{r}$ which can be written in Cartesian and spherical coordinates as follows

$$(x,\ y,\ z) \longrightarrow (x',\ y',\ z') = (-x,\ -y,\ -z), \tag{5.192}$$

$$(r,\ \theta,\ \varphi) \longrightarrow (r',\ \theta',\ \varphi') = (r,\ \pi-\theta,\ \varphi+\pi). \tag{5.193}$$

We can establish a connection between the spherical harmonics and the Legendre polynomials by simply taking $m = 0$. Then equation (5.187) yields

$$Y_{l0}(\theta,\varphi) = \frac{(-1)^l}{2^l l!}\sqrt{\frac{2l+1}{4\pi}}\frac{d^l}{d(\cos\theta)^l}(\sin\theta)^{2l} = \sqrt{\frac{2l+1}{4\pi}}P_l(\cos\theta), \tag{5.194}$$

with

$$P_l(\cos\theta) = \frac{1}{2^l l!}\frac{d^l}{d(\cos\theta)^l}(\cos^2\theta - 1)^l. \tag{5.195}$$

From the expression of Y_{lm}, we can verify that

$$Y_{lm}(0,\varphi) = \sqrt{\frac{2l+1}{4\pi}}\delta_{m,0}. \tag{5.196}$$

The expressions for the spherical harmonics corresponding to $l = 0$, $l = 1$, and $l = 2$ are listed in Table 5.2.

Spherical harmonics in Cartesian coordinates

Note that $Y_{lm}(\theta,\varphi)$ can also be expressed in terms of the Cartesian coordinates. For this, we need only to substitute

$$\sin\theta\cos\varphi = \frac{x}{r}, \qquad \sin\theta\sin\varphi = \frac{y}{r}, \qquad \cos\theta = \frac{z}{r} \tag{5.197}$$

in the expression for $Y_{lm}(\theta,\varphi)$.

As an illustration, let us show how to derive the Cartesian expressions for Y_{10} and $Y_{1,\pm1}$. Substituting $\cos\theta = z/r$ into $Y_{10}(\theta,\varphi) = \sqrt{3/4\pi}\cos\theta\, Y_{10}$, we have

$$Y_{10}(x,y,z) = \sqrt{\frac{3}{4\pi}}\frac{z}{r} = \sqrt{\frac{3}{4\pi}}\frac{z}{\sqrt{x^2+y^2+z^2}}. \tag{5.198}$$

Table 5.2: Spherical harmonics and their expressions in Cartesian coordinates.

$Y_{lm}(\theta,\varphi)$	$Y_{lm}(x,y,z)$
$Y_{00}(\theta,\varphi)=\frac{1}{\sqrt{4\pi}}$	$Y_{00}(x,y,z)=\frac{1}{\sqrt{4\pi}}$
$Y_{10}(\theta,\varphi)=\sqrt{\frac{3}{4\pi}}\cos\theta$	$Y_{10}(x,y,z)=\sqrt{\frac{3}{4\pi}}\frac{z}{r}$
$Y_{1,\pm1}(\theta,\varphi)=\mp\sqrt{\frac{3}{8\pi}}e^{\pm i\varphi}\sin\theta$	$Y_{1,\pm1}(x,y,z)=\mp\sqrt{\frac{3}{8\pi}}\frac{x\pm iy}{r}$
$Y_{20}(\theta,\varphi)=\sqrt{\frac{5}{16\pi}}(3\cos^2\theta-1)$	$Y_{20}(x,y,z)=\sqrt{\frac{5}{16\pi}}\frac{3z^2-r^2}{r^2}$
$Y_{2,\pm1}(\theta,\varphi)=\mp\sqrt{\frac{15}{8\pi}}e^{\pm i\varphi}\sin\theta\cos\theta$	$Y_{2,\pm1}(x,y,z)=\mp\sqrt{\frac{15}{8\pi}}\frac{(x\pm iy)z}{r^2}$
$Y_{2,\pm2}(\theta,\varphi)=\sqrt{\frac{15}{32\pi}}e^{\pm 2i\varphi}\sin^2\theta$	$Y_{2,\pm2}(x,y,z)=\sqrt{\frac{15}{32\pi}}\frac{x^2-y^2\pm 2ixy}{r^2}$

Using $\sin\theta\cos\varphi = x/r$ and $\sin\theta\sin\varphi = y/r$, we obtain

$$\frac{x\pm iy}{r}=\sin\theta\cos\varphi\pm i\sin\theta\sin\varphi=\sin\theta\, e^{\pm i\varphi}, \tag{5.199}$$

which, when substituted into $Y_{1\pm1}(\theta,\varphi)=\mp\sqrt{3/8\pi}\sin\theta\, e^{\pm i\varphi}$, leads to

$$Y_{1,\pm1}(x,y,z)=\mp\sqrt{\frac{3}{8\pi}}\frac{x\pm iy}{r}. \tag{5.200}$$

Following the same procedure, we can derive the Cartesian expressions of the remaining harmonics; for a listing, see Table 5.2.

Example 5.5 (Application of ladder operators to spherical harmonics)

(a) Use the relation $Y_{l0}(\theta,\varphi)=\sqrt{(2l+1)/4\pi}P_l(\cos\theta)$ to find the expression of $Y_{30}(\theta,\varphi)$.

(b) Find the expression of Y_{30} in Cartesian coordinates.

(c) Use the expression of $Y_{30}(\theta,\varphi)$ to infer those of $Y_{3,\pm1}(\theta,\varphi)$.

Solution

(a) From Table 5.1 we have $P_3(\cos\theta)=\frac{1}{2}(5\cos^3\theta-3\cos\theta)$; hence

$$Y_{30}(\theta,\varphi)=\sqrt{\frac{7}{4\pi}}P_3(\cos\theta)=\sqrt{\frac{7}{16\pi}}(5\cos^3\theta-3\cos\theta). \tag{5.201}$$

(b) Since $\cos\theta = z/r$, we have $5\cos^3\theta-3\cos\theta = 5\cos\theta(5\cos^2\theta-3) = z(5z^2-3r^2)/r^3$; hence

$$Y_{30}(x,y,z)=\sqrt{\frac{7}{16\pi}}\frac{z}{r^3}(5z^2-3r^2). \tag{5.202}$$

(c) To find Y_{31} from Y_{30}, we need to apply the ladder operator $\hat{L}_+$ on Y_{30} in two ways: first, algebraically

$$\hat{L}_+ Y_{30} = \hbar\sqrt{3(3+1)-0}\, Y_{31} = 2\hbar\sqrt{3}\, Y_{31} \tag{5.203}$$

and hence

$$Y_{31} = \frac{1}{2\hbar\sqrt{3}}\hat{L}_+ Y_{30}; \tag{5.204}$$

then we use the differential form (5.141) of $\hat{L}_+$:

$$\begin{aligned}\hat{L}_+ Y_{30}(\theta,\varphi) &= \hbar e^{i\varphi}\left[\frac{\partial}{\partial\theta} + i\frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\varphi}\right] Y_{30}(\theta,\varphi)\\ &= \hbar\sqrt{\frac{7}{16\pi}}e^{i\varphi}\left[\frac{\partial}{\partial\theta} + i\frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\varphi}\right](5\cos^3\theta - 3\cos\theta)\\ &= -3\hbar\sqrt{\frac{7}{16\pi}}\sin\theta(5\cos^2\theta - 1)e^{i\varphi}.\end{aligned} \tag{5.205}$$

Inserting (5.205) into (5.204) we end up with

$$Y_{31} = \frac{1}{2\hbar\sqrt{3}}\hat{L}_+ Y_{30} = -\sqrt{\frac{21}{64\pi}}\sin\theta(5\cos^2\theta - 1)e^{i\varphi}. \tag{5.206}$$

Now, to find $Y_{3,-1}$ from Y_{30}, we also need to apply $\hat{L}_-$ on Y_{30} in two ways:

$$\hat{L}_- Y_{30} = \hbar\sqrt{3(3+1)-0}\, Y_{3,-1} = 2\hbar\sqrt{3} Y_{3,-1} \tag{5.207}$$

and hence

$$Y_{3,-1} = \frac{1}{2\hbar\sqrt{3}}\hat{L}_- Y_{30}; \tag{5.208}$$

then we use the differential form (5.141) of $\hat{L}_-$:

$$\begin{aligned}\hat{L}_- Y_{30}(\theta,\varphi) &= -\hbar e^{-i\varphi}\left[\frac{\partial}{\partial\theta} - i\frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\varphi}\right] Y_{30}(\theta,\varphi)\\ &= -\hbar\sqrt{\frac{7}{16\pi}}e^{-i\varphi}\left[\frac{\partial}{\partial\theta} - i\frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\varphi}\right](5\cos^3\theta - 3\cos\theta)\\ &= 3\hbar\sqrt{\frac{7}{16\pi}}\sin\theta(5\cos^2\theta - 1)e^{-i\varphi}.\end{aligned} \tag{5.209}$$

Inserting (5.209) into (5.208), we obtain

$$Y_{3,-1} = \frac{1}{2\hbar\sqrt{3}}\hat{L}_- Y_{30} = \sqrt{\frac{21}{64\pi}}\sin\theta(5\cos^2\theta - 1)e^{-i\varphi}. \tag{5.210}$$

5.8 Solved Problems

Problem 5.1

(a) Show that $\Delta J_x \Delta J_y = \hbar^2[j(j+1) - m^2]/2$, where $\Delta J_x = \sqrt{\langle \hat{J}_x^2 \rangle - \langle \hat{J}_x \rangle^2}$ and the same for ΔJ_y.

(b) Show that this relation is consistent with $\Delta J_x \Delta J_y \geq (\hbar/2)\,|\langle \hat{J}_z \rangle| = \hbar^2 m/2$.

Solution

(a) First, note that $\langle \hat{J}_x \rangle$ and $\langle \hat{J}_y \rangle$ are zero, since

$$\langle \hat{J}_x \rangle = \frac{1}{2}\langle j,\, m \mid \hat{J}_+ \mid j,\, m \rangle + \frac{1}{2}\langle j,\, m \mid \hat{J}_- \mid j,\, m \rangle = 0. \tag{5.211}$$

As for $\langle \hat{J}_x^2 \rangle$ and $\langle \hat{J}_y^2 \rangle$, they are given by

$$\langle \hat{J}_x^2 \rangle = \frac{1}{4}\langle (\hat{J}_+ + \hat{J}_-)^2 \rangle = \frac{1}{4}\langle \hat{J}_+^2 + \hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+ + \hat{J}_-^2 \rangle, \tag{5.212}$$

$$\langle \hat{J}_y^2 \rangle = \frac{1}{4}\langle (\hat{J}_+ - \hat{J}_-)^2 \rangle = -\frac{1}{4}\langle \hat{J}_+^2 - \hat{J}_+\hat{J}_- - \hat{J}_-\hat{J}_+ + \hat{J}_-^2 \rangle. \tag{5.213}$$

Since $\langle \hat{J}_+^2 \rangle = \langle \hat{J}_-^2 \rangle = 0$, we see that

$$\langle \hat{J}_x^2 \rangle = \frac{1}{4}\langle \hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+ \rangle = \langle \hat{J}_y^2 \rangle. \tag{5.214}$$

Using the fact that

$$\langle \hat{J}_x^2 \rangle + \langle \hat{J}_y^2 \rangle = \langle \hat{\vec{J}}^2 \rangle - \langle \hat{J}_z^2 \rangle \tag{5.215}$$

along with $\langle \hat{J}_x^2 \rangle = \langle \hat{J}_y^2 \rangle$, we see that

$$\langle \hat{J}_x^2 \rangle = \langle \hat{J}_y^2 \rangle = \frac{1}{2}[\langle \hat{\vec{J}}^2 \rangle - \langle \hat{J}_z^2 \rangle]. \tag{5.216}$$

Now, since $|\, j,\, m\rangle$ is a joint eigenstate of $\hat{\vec{J}}^2$ and $\hat{J}_z$ with eigenvalues $j(j+1)\hbar^2$ and $m\hbar$, we can easily see that the expressions of $\langle \hat{J}_x^2 \rangle$ and $\langle \hat{J}_y^2 \rangle$ are given by

$$\boxed{\langle \hat{J}_x^2 \rangle = \langle \hat{J}_y^2 \rangle = \frac{1}{2}[\langle \hat{\vec{J}}^2 \rangle - \langle \hat{J}_z^2 \rangle] = \frac{\hbar^2}{2}\left[j(j+1) - m^2\right].} \tag{5.217}$$

Hence $\Delta J_x \Delta J_y$ is given by

$$\Delta J_x \Delta J_y = \sqrt{\langle \hat{J}_x^2 \rangle \langle \hat{J}_y^2 \rangle} = \frac{\hbar^2}{2}[j(j+1) - m^2]. \tag{5.218}$$

(b) Since $j \geq m$ (because $m = -j, -j+1, \ldots, j-1, j$), we have

$$j(j+1) - m^2 \geq m(m+1) - m^2 = m, \tag{5.219}$$

from which we infer that $\Delta J_x \Delta J_y \geq \hbar^2 m/2$, or

$$\Delta J_x \Delta J_y \geq \frac{\hbar}{2}|\langle \hat{J}_z \rangle|. \tag{5.220}$$

Problem 5.2
Find the energy levels of a particle which is free except that it is constrained to move on the surface of a sphere of radius r.

Solution
This system consists of a particle that is constrained to move on the surface of a sphere but free from the influence of any other potential; it is called a *rigid rotor*. Since $V = 0$ the energy of this system is purely kinetic; the Hamiltonian of the rotor is

$$\hat{H} = \frac{\hat{\vec{L}}^2}{2I}, \tag{5.221}$$

where $I = mr^2$ is the moment of inertia of the particle with respect to the origin. In deriving this relation, we have used the fact that $H = p^2/2m = (rp)^2/2mr^2 = L^2/2I$, since $L = |\vec{r} \times \vec{p}| = rp$.

The wave function of the system is clearly independent of the radial degree of freedom, for it is constant. The Schrödinger equation is thus given by

$$\hat{H}\psi(\theta, \varphi) = \frac{\hat{\vec{L}}^2}{2I}\psi(\theta, \varphi) = E\psi(\theta, \varphi). \tag{5.222}$$

Since the eigenstates of $\hat{\vec{L}}^2$ are the spherical harmonics $Y_{lm}(\theta, \varphi)$, the corresponding energy eigenvalues are given by

$$E_l = \frac{\hbar^2}{2I} l(l+1), \qquad l = 0, 1, 2, 3, \ldots, \tag{5.223}$$

and the Schrödinger equation by

$$\frac{\hat{\vec{L}}^2}{2I} Y_{lm}(\theta, \varphi) = \frac{\hbar^2}{2I} l(l+1) Y_{lm}(\theta, \varphi). \tag{5.224}$$

Note that the energy levels do not depend on the azimuthal quantum number m. This means that there are $(2l + 1)$ eigenfunctions Y_{l-l}, Y_{l-l+1}, $\ldots$, Y_{ll-1}, Y_{ll} corresponding to the same energy. Thus, every energy level E_l is $(2l + 1)$-fold degenerate. This is due to the fact that the rotor's Hamiltonian, $\hat{\vec{L}}^2/2I$, commutes with $\hat{\vec{L}}$. That is, the Hamiltonian is independent of the orientation of $\hat{\vec{L}}$ in space; hence the energy spectrum does not depend on the component of $\hat{\vec{L}}$ in any particular direction.

Problem 5.3
Find the rotational energy levels of a diatomic molecule.

Solution
Consider two molecules of masses m_1 and m_2 separated by a constant distance $\vec{r}$. Let r_1 and r_2 be their distances from the center of mass, i.e., $m_1 r_1 = m_2 r_2$. The moment of inertia of the diatomic molecule is

$$I = m_1 r_1^2 + m_2 r_2^2 \equiv \mu r^2, \tag{5.225}$$

where $r = |\vec{r}_1 - \vec{r}_2|$ and where μ is their reduced mass, $\mu = m_1 m_2/(m_1 + m_2)$. The total angular momentum is given by

$$|\hat{\vec{L}}| = m_1 r_1 r_1 \omega + m_2 r_2 r_2 \omega = I\omega = \mu r^2 \omega \tag{5.226}$$

and the Hamiltonian by

$$\hat{H} = \frac{\hat{\vec{L}}^2}{2I} = \frac{\hat{\vec{L}}^2}{2\mu r^2}. \tag{5.227}$$

The corresponding eigenvalue equation

$$\hat{H}\,|l,\, m\rangle = \frac{\hat{\vec{L}}^2}{2\mu r^2}\,|l,\, m\rangle = \frac{l(l+1)\hbar^2}{2\mu r^2}\,|l,\, m\rangle, \tag{5.228}$$

shows that the eigenenergies are $(2l + 1)$-fold degenerate and given by

$$E_l = \frac{l(l+1)\hbar^2}{2\mu r^2}. \tag{5.229}$$

Problem 5.4

(a) Find the eigenvalues and eigenstates of the spin operator $\vec{S}$ of an electron in the direction of a unit vector $\vec{n}$; assume that $\vec{n}$ lies in the xz plane.

(b) Find the probability of measuring $\hat{S}_z = +\hbar/2$.

Solution

(a) In this question we want to solve

$$\vec{n} \cdot \vec{S}|\lambda\rangle = \frac{\hbar}{2}\lambda|\lambda\rangle, \tag{5.230}$$

where $\vec{n}$ is given by $\vec{n} = (\sin\theta\,\hat{\imath} + \cos\theta\,\hat{k})$, because it lies in the xz plane, with $0 \le \theta \le \pi$. We can thus write

$$\vec{n} \cdot \vec{S} = (\sin\theta\,\hat{\imath} + \cos\theta\,\hat{k}) \cdot (S_x\hat{\imath} + S_y\hat{\jmath} + S_z\hat{k}) = S_x \sin\theta + S_z \cos\theta. \tag{5.231}$$

Using the spin matrices

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{5.232}$$

we can write (5.231) in the following matrix form:

$$\vec{n} \cdot \vec{S} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\sin\theta + \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\cos\theta = \frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}. \tag{5.233}$$

The diagonalization of this matrix leads to the following secular equation:

$$-\frac{\hbar^2}{4}(\cos\theta - \lambda)(\cos\theta + \lambda) - \frac{\hbar^2}{4}\sin^2\theta = 0, \tag{5.234}$$

which in turn leads as expected to the eigenvalues $\lambda = \pm 1$.

The eigenvector corresponding to $\lambda = 1$ can be obtained from

$$\frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} a \\ b \end{pmatrix}. \tag{5.235}$$

This matrix equation can be reduced to a single equation

$$a \sin\frac{1}{2}\theta = b\cos\frac{1}{2}\theta. \tag{5.236}$$

Combining this equation with the normalization condition $|a|^2 + |b|^2 = 1$, we infer that $a = \cos\frac{1}{2}\theta$ and $b = \sin\frac{1}{2}\theta$; hence the eigenvector corresponding to $\lambda = 1$ is

$$|\lambda_+\rangle = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}. \tag{5.237}$$

Proceeding in the same way, we can easily obtain the eigenvector for $\lambda = -1$:

$$|\lambda_-\rangle = \begin{pmatrix} -\sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix}. \tag{5.238}$$

(b) Let us write $|\lambda_\pm\rangle$ of (5.237) and (5.238) in terms of the spin-up and spin-down eigenvectors, $\left|\frac{1}{2}, \frac{1}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$:

$$|\lambda_+\rangle = \cos\frac{1}{2}\theta\left|\frac{1}{2}, \frac{1}{2}\right\rangle + \sin\frac{1}{2}\theta\left|\frac{1}{2}, -\frac{1}{2}\right\rangle, \tag{5.239}$$

$$|\lambda_-\rangle = -\sin\frac{1}{2}\theta\left|\frac{1}{2}, \frac{1}{2}\right\rangle + \cos\frac{1}{2}\theta\left|\frac{1}{2}, -\frac{1}{2}\right\rangle. \tag{5.240}$$

We see that the probability of measuring $\hat{S}_z = +\hbar/2$ is given by

$$\left|\left\langle\frac{1}{2}, \frac{1}{2}\middle|\lambda_+\right\rangle\right|^2 = \cos^2\frac{1}{2}\theta. \tag{5.241}$$

Problem 5.5

(a) Find the eigenvalues and eigenstates of the spin operator $\vec{S}$ of an electron in the direction of a unit vector $\vec{n}$, where $\vec{n}$ is *arbitrary*.

(b) Find the probability of measuring $\hat{S}_z = -\hbar/2$.

(c) Assuming that the eigenvectors of the spin calculated in (a) correspond to $t = 0$, find these eigenvectors at time t.

Solution

(a) We need to solve

$$\vec{n}\cdot\vec{S}|\lambda\rangle = \frac{\hbar}{2}\lambda|\lambda\rangle, \tag{5.242}$$

where $\vec{n}$, a unit vector pointing along an *arbitrary* direction, is given in spherical coordinates by

$$\vec{n} = (\sin\theta\cos\varphi)\hat{i} + (\sin\theta\sin\varphi)\hat{j} + (\cos\theta)\hat{k}, \tag{5.243}$$

with $0 \le \theta \le \pi$ and $0 \le \varphi \le 2\pi$. We can thus write

$$\begin{aligned}\vec{n}\cdot\vec{S} &= (\sin\theta\cos\varphi\,\hat{i} + \sin\theta\sin\varphi\,\hat{j} + \cos\theta\,\hat{k})\cdot(S_x\hat{i} + S_y\hat{j} + S_z\hat{k}) \\ &= S_x\sin\theta\cos\varphi + S_y\sin\theta\sin\varphi + S_z\cos\theta.\end{aligned} \tag{5.244}$$

Using the spin matrices, we can write this equation in the following matrix form:

$$\begin{aligned}\vec{n}\cdot\vec{S} &= \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\sin\theta\cos\varphi + \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\sin\theta\sin\varphi + \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\cos\theta \\ &= \frac{\hbar}{2}\begin{pmatrix}\cos\theta & \sin\theta(\cos\varphi - i\sin\varphi)\\ \sin\theta(\cos\varphi + i\sin\varphi) & -\cos\theta\end{pmatrix} \\ &= \frac{\hbar}{2}\begin{pmatrix}\cos\theta & e^{-i\varphi}\sin\theta\\ e^{i\varphi}\sin\theta & -\cos\theta\end{pmatrix}.\end{aligned} \tag{5.245}$$

Diagonalization of this matrix leads to the secular equation

$$-\frac{\hbar^2}{4}(\cos\theta - \lambda)(\cos\theta + \lambda) - \frac{\hbar^2}{4}\sin^2\theta = 0, \tag{5.246}$$

which in turn leads to the eigenvalues $\lambda = \pm 1$.

The eigenvector corresponding to $\lambda = 1$ can be obtained from

$$\frac{\hbar}{2}\begin{pmatrix}\cos\theta & e^{-i\varphi}\sin\theta\\ e^{i\varphi}\sin\theta & -\cos\theta\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}a\\ b\end{pmatrix}, \tag{5.247}$$

which leads to

$$a\cos\theta + be^{-i\varphi}\sin\theta = a \tag{5.248}$$

or

$$a(1 - \cos\theta) = be^{-i\varphi}\sin\theta. \tag{5.249}$$

Using the relations $1 - \cos\theta = 2\sin^2\frac{1}{2}\theta$ and $\sin\theta = 2\cos\frac{1}{2}\theta\sin\frac{1}{2}\theta$, we have

$$b = a\tan\frac{1}{2}\theta e^{i\varphi}. \tag{5.250}$$

Combining this equation with the normalization condition $|a|^2 + |b|^2 = 1$, we obtain $a = \cos\frac{1}{2}\theta$ and $b = e^{i\varphi}\sin\frac{1}{2}\theta$. Thus, the eigenvector corresponding to $\lambda = 1$ is

$$|\lambda_+\rangle = \begin{pmatrix}\cos(\theta/2)\\ e^{i\varphi}\sin(\theta/2)\end{pmatrix}. \tag{5.251}$$

A similar treatment leads to the eigenvector for $\lambda = -1$:

$$|\lambda_-\rangle = \begin{pmatrix}-\sin(\theta/2)\\ e^{i\varphi}\cos(\theta/2)\end{pmatrix}. \tag{5.252}$$

(b) Write $|\lambda_-\rangle$ of (5.252) in terms of $\left|\frac{1}{2}, \frac{1}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$:

$$|\lambda_+\rangle = \cos\frac{1}{2}\theta\left|\frac{1}{2}, \frac{1}{2}\right\rangle + e^{i\varphi}\sin\frac{1}{2}\theta\left|\frac{1}{2}, -\frac{1}{2}\right\rangle, \tag{5.253}$$

$$|\lambda_-\rangle = -\sin\frac{1}{2}\theta\left|\frac{1}{2}, \frac{1}{2}\right\rangle + e^{i\varphi}\cos\frac{1}{2}\theta\left|\frac{1}{2}, -\frac{1}{2}\right\rangle. \tag{5.254}$$

We can then obtain the probability of measuring $\hat{S}_z = -\hbar/2$:

$$\left|\left\langle\frac{1}{2}, -\frac{1}{2}\right|\lambda_-\right\rangle\Big|^2 = \cos^2\frac{1}{2}\theta. \tag{5.255}$$

(c) The spin's eigenstates at time t are given by

$$|\lambda_+(t)\rangle = e^{-iE_+t/\hbar}\cos\frac{1}{2}\theta\left|\frac{1}{2}, \frac{1}{2}\right\rangle + e^{i(\varphi - E_-t/\hbar)}\sin\frac{1}{2}\theta\left|\frac{1}{2}, -\frac{1}{2}\right\rangle, \tag{5.256}$$

$$|\lambda_-(t)\rangle = -e^{-iE_+t/\hbar}\sin\frac{1}{2}\theta\left|\frac{1}{2}, \frac{1}{2}\right\rangle + e^{i(\varphi - E_-t/\hbar)}\cos\frac{1}{2}\theta\left|\frac{1}{2}, -\frac{1}{2}\right\rangle, \tag{5.257}$$

where $E_\pm$ are the energy eigenvalues corresponding to the spin-up and spin-down states, respectively.

Problem 5.6

The Hamiltonian of a system is $\hat{H} = \varepsilon\vec{\sigma}\cdot\vec{n}$, where ε is a constant having the dimensions of energy, $\vec{n}$ is an arbitrary unit vector, and σ_x, σ_y, and σ_z are the Pauli matrices.

(a) Find the energy eigenvalues and normalized eigenvectors of $\hat{H}$.

(b) Find a transformation matrix that diagonalizes $\hat{H}$.

Solution

(a) Using the Pauli matrices $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and the expression of an arbitrary unit vector in spherical coordinates $\vec{n} = (\sin\theta\cos\varphi)\hat{\imath} + (\sin\theta\sin\varphi)\hat{\jmath} + (\cos\theta)\hat{k}$, we can rewrite the Hamiltonian

$$\hat{H} = \varepsilon\vec{\sigma}\cdot\vec{n} = \varepsilon\left(\sigma_x\sin\theta\cos\varphi + \sigma_y\sin\theta\sin\varphi + \sigma_z\cos\theta\right) \tag{5.258}$$

in the following matrix form:

$$\hat{H} = \varepsilon\begin{pmatrix} \cos\theta & \exp(-i\varphi)\sin\theta \\ \exp(i\varphi)\sin\theta & -\cos\theta \end{pmatrix}. \tag{5.259}$$

The eigenvalues of $\hat{H}$ are obtained by solving the secular equation $\det(H - E) = 0$, or

$$(\varepsilon\cos\theta - E)(-\varepsilon\cos\theta - E) - \varepsilon^2\sin^2\theta = 0, \tag{5.260}$$

which yields two eigenenergies $E_1 = \varepsilon$ and $E_2 = -\varepsilon$.

The energy eigenfunctions are obtained from

$$\varepsilon\begin{pmatrix} \cos\theta & \exp(-i\varphi)\sin\theta \\ \exp(i\varphi)\sin\theta & -\cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = E\begin{pmatrix} x \\ y \end{pmatrix}. \tag{5.261}$$

For the case $E = E_1 = \varepsilon$, this equation yields

$$(\cos\theta - 1)x + y\sin\theta\exp(-i\varphi) = 0, \tag{5.262}$$

which in turn leads to

$$\frac{x}{y} = \frac{\sin\theta\exp(-i\varphi)}{1-\cos\theta} = \frac{\cos\theta/2\exp(-i\varphi/2)}{\sin\theta/2\exp(i\varphi/2)}; \tag{5.263}$$

hence

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \exp(-i\varphi/2)\cos(\theta/2) \\ \exp(i\varphi/2)\sin(\theta/2) \end{pmatrix}; \tag{5.264}$$

this vector is normalized. Similarly, in the case where $E = E_2 = -\varepsilon$, we can show that the second normalized eigenvector is

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} -\exp(-i\varphi/2)\sin(\theta/2) \\ \exp(i\varphi/2)\cos(\theta/2) \end{pmatrix}. \tag{5.265}$$

(b) A transformation $\hat{U}$ that diagonalizes $\hat{H}$ can be obtained from the two eigenvectors obtained in part (a): $U_{11} = x_1$, $U_{21} = y_1$, $U_{12} = x_2$, $U_{22} = y_2$. That is,

$$U = \begin{pmatrix} \exp(-i\varphi/2)\cos(\theta/2) & -\exp(-i\varphi/2)\sin(\theta/2) \\ \exp(i\varphi/2)\sin(\theta/2) & \exp(i\varphi/2)\cos(\theta/2) \end{pmatrix}. \tag{5.266}$$

Note that this matrix is unitary, since $U^\dagger = U^{-1}$ and $\det(U) = 1$. We can ascertain that

$$\hat{U}\hat{H}\hat{U}^\dagger = \begin{pmatrix} \varepsilon & 0 \\ 0 & -\varepsilon \end{pmatrix}. \tag{5.267}$$

Problem 5.7

Consider a system of total angular momentum $j = 1$. As shown in (5.73) and (5.75), the operators $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$ are given by

$$\hat{J}_x = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \hat{J}_y = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \hat{J}_z = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{5.268}$$

(a) What are the possible values when measuring $\hat{J}_x$?
(b) Calculate $\langle\hat{J}_z\rangle$, $\langle\hat{J}_z^2\rangle$, and ΔJ_z if the system is in the state $j_x = -\hbar$.
(c) Repeat (b) for $\langle\hat{J}_y\rangle$, $\langle\hat{J}_y^2\rangle$, and ΔJ_y.
(d) If the system were initially in state $|\psi\rangle = \frac{1}{\sqrt{14}}\begin{pmatrix} -\sqrt{3} \\ 2\sqrt{2} \\ \sqrt{3} \end{pmatrix}$, what values will one obtain when measuring $\hat{J}_x$ and with what probabilities?

Solution

(a) According to Postulate 2 of Chapter 3, the results of the measurements are given by the eigenvalues of the measured quantity. Here the eigenvalues of $\hat{J}_x$, which are obtained by diagonalizing the matrix J_x, are $j_x = -\hbar, 0$, and $\hbar$; the respective (normalized) eigenstates are

$$|-1\rangle = \frac{1}{2}\begin{pmatrix} -1 \\ \sqrt{2} \\ -1 \end{pmatrix}, \qquad |0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \qquad |1\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}. \tag{5.269}$$

(b) If the system is in the state $j_x = -\hbar$, its eigenstate is given by $|-1\rangle$. In this case $\langle \hat{J}_z \rangle$ and $\langle \hat{J}_z^2 \rangle$ are given by

$$\langle -1|\hat{J}_z|-1\rangle = \frac{\hbar}{4}\begin{pmatrix} -1 & \sqrt{2} & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} -1 \\ \sqrt{2} \\ -1 \end{pmatrix} = 0, \tag{5.270}$$

$$\langle -1|\hat{J}_z^2|-1\rangle = \frac{\hbar^2}{4}\begin{pmatrix} -1 & \sqrt{2} & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ \sqrt{2} \\ -1 \end{pmatrix} = \frac{\hbar^2}{2}. \tag{5.271}$$

Thus, the uncertainty ΔJ_z is given by

$$\Delta J_z = \sqrt{\langle -1|\hat{J}_z^2|-1\rangle - \langle -1|\hat{J}_z|-1\rangle^2} = \sqrt{\frac{\hbar^2}{2}} = \frac{\hbar}{\sqrt{2}}. \tag{5.272}$$

(c) Following the same procedure in (b), we have

$$\langle -1|\hat{J}_y|-1\rangle = \frac{\hbar}{4\sqrt{2}}\begin{pmatrix} -1 & \sqrt{2} & -1 \end{pmatrix}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}\begin{pmatrix} -1 \\ \sqrt{2} \\ -1 \end{pmatrix} = 0, \tag{5.273}$$

$$\langle -1|\hat{J}_y^2|-1\rangle = \frac{\hbar^2}{8}\begin{pmatrix} -1 & \sqrt{2} & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ \sqrt{2} \\ -1 \end{pmatrix} = \frac{\hbar^2}{2}; \tag{5.274}$$

hence

$$\Delta J_y = \sqrt{\langle -1|\hat{J}_y^2|-1\rangle - \langle -1|\hat{J}_y|-1\rangle^2} = \frac{\hbar}{\sqrt{2}}. \tag{5.275}$$

(d) We can express $|\psi\rangle = \frac{1}{\sqrt{14}}\begin{pmatrix} -\sqrt{3} \\ 2\sqrt{2} \\ \sqrt{3} \end{pmatrix}$ in terms of the eigenstates (5.269) as

$$\frac{1}{\sqrt{14}}\begin{pmatrix} -\sqrt{3} \\ 2\sqrt{2} \\ \sqrt{3} \end{pmatrix} = \sqrt{\frac{2}{7}}\frac{1}{2}\begin{pmatrix} -1 \\ \sqrt{2} \\ -1 \end{pmatrix} + \sqrt{\frac{3}{7}}\frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \sqrt{\frac{2}{7}}\frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} \tag{5.276}$$

or

$$|\psi\rangle = \sqrt{\frac{2}{7}}|-1\rangle + \sqrt{\frac{3}{7}}|0\rangle + \sqrt{\frac{2}{7}}|1\rangle. \tag{5.277}$$

A measurement of $\hat{J}_x$ on a system initially in the state (5.277) yields a value $j_x = -\hbar$ with probability

$$P_{-1} = |\langle -1 \mid \psi\rangle|^2 = \left|\sqrt{\frac{2}{7}}\langle -1|-1\rangle + \sqrt{\frac{3}{7}}\langle -1|0\rangle + \sqrt{\frac{2}{7}}\langle -1|1\rangle\right|^2 = \frac{2}{7}, \tag{5.278}$$

since $\langle -1|0\rangle = \langle -1|1\rangle = 0$ and $\langle -1|-1\rangle = 1$, and the values $j_x = 0$ and $j_x = \hbar$ with the respective probabilities

$$P_0 = |\langle 0 \mid \psi\rangle|^2 = \left|\sqrt{\frac{3}{7}}\langle 0|0\rangle\right|^2 = \frac{3}{7}, \qquad P_1 = |\langle 1 \mid \psi\rangle|^2 = \left|\sqrt{\frac{2}{7}}\langle 1|1\rangle\right|^2 = \frac{2}{7}. \tag{5.279}$$

Problem 5.8
Consider a particle of total angular momentum $j = 1$. Find the matrix for the component of $\vec{J}$ along a unit vector with arbitrary direction $\vec{n}$. Find its eigenvalues and eigenvectors.

Solution
Since $\vec{J} = J_x\hat{\imath} + J_y\hat{\jmath} + J_z\hat{k}$ and $\vec{n} = (\sin\theta\cos\varphi)\hat{\imath} + (\sin\theta\sin\varphi)\hat{\jmath} + (\cos\theta)\hat{k}$, the component of $\vec{J}$ along $\vec{n}$ is

$$\vec{n}\cdot\vec{J} = J_x\sin\theta\cos\varphi + J_y\sin\theta\sin\varphi + J_z\cos\theta, \tag{5.280}$$

with $0 \le \theta \le \pi$ and $0 \le \varphi \le 2\pi$; the matrices of $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$ are given by (5.268). We can therefore write this equation in the following matrix form:

$$\begin{aligned}\vec{n}\cdot\vec{J} &= \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\sin\theta\cos\varphi + \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}\sin\theta\sin\varphi \\ &+ \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}\cos\theta = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} \sqrt{2}\cos\theta & e^{-i\varphi}\sin\theta & 0 \\ e^{i\varphi}\sin\theta & 0 & e^{-i\varphi}\sin\theta \\ 0 & e^{i\varphi}\sin\theta & -\sqrt{2}\cos\theta \end{pmatrix}. \end{aligned} \tag{5.281}$$

The diagonalization of this matrix leads to the eigenvalues $\lambda_1 = -\hbar$, $\lambda_2 = 0$, and $\lambda_3 = \hbar$; the corresponding eigenvectors are given by

$$|\lambda_1\rangle = \frac{1}{2}\begin{pmatrix} (1-\cos\theta)e^{-i\varphi} \\ -\frac{2}{\sqrt{2}}\sin\theta \\ (1+\cos\theta)e^{i\varphi} \end{pmatrix}, \qquad |\lambda_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -e^{-i\varphi}\sin\theta \\ \sqrt{2}\cos\theta \\ e^{i\varphi}\sin\theta \end{pmatrix}, \tag{5.282}$$

$$|\lambda_3\rangle = \frac{1}{2}\begin{pmatrix} (1+\cos\theta)e^{-i\varphi} \\ \frac{2}{\sqrt{2}}\sin\theta \\ (1-\cos\theta)e^{i\varphi} \end{pmatrix}. \tag{5.283}$$

Problem 5.9
Consider a system which is initially in the state

$$\psi(\theta, \varphi) = \frac{1}{\sqrt{5}} Y_{1,-1}(\theta, \varphi) + \sqrt{\frac{3}{5}} Y_{10}(\theta, \varphi) + \frac{1}{\sqrt{5}} Y_{11}(\theta, \varphi).$$

(a) Find $\langle \psi \mid \hat{L}_+ \mid \psi \rangle$.
(b) If $\hat{L}_z$ were measured what values will one obtain and with what probabilities?
(c) If after measuring $\hat{L}_z$ we find $l_z = -\hbar$, calculate the uncertainties ΔL_x and ΔL_y and their product $\Delta L_x \Delta L_y$.

Solution

(a) Let us use a lighter notation for $\mid \psi \rangle$: $\mid \psi \rangle = \frac{1}{\sqrt{5}} \mid 1, -1 \rangle + \sqrt{\frac{3}{5}} \mid 1, 0 \rangle + \frac{1}{\sqrt{5}} \mid 1, 1 \rangle$. From (5.56) we can write $\hat{L}_+ \mid l, m \rangle = \hbar \sqrt{l(l+1) - m(m+1)} \mid l, m+1 \rangle$; hence the only terms that survive in $\langle \psi \mid \hat{L}_+ \mid \psi \rangle$ are

$$\langle \psi \mid \hat{L}_+ \mid \psi \rangle = \frac{\sqrt{3}}{5} \langle 1, 0 | \hat{L}_+ \mid 1, -1 \rangle + \frac{\sqrt{3}}{5} \langle 1, 1 | \hat{L}_+ \mid 1, 0 \rangle = \frac{2\sqrt{6}}{5} \hbar, \tag{5.284}$$

since $\langle 1, 0 | \hat{L}_+ \mid 1, -1 \rangle = \langle 1, 1 | \hat{L}_+ \mid 1, 0 \rangle = \sqrt{2}\hbar$.

(b) If $\hat{L}_z$ were measured, we will find three values $l_z = -\hbar$, 0, and $\hbar$. The probability of finding the value $l_z = -\hbar$ is

$$P_{-1} = \left| \langle 1, -1 \mid \psi \rangle \right|^2 = \left| \frac{1}{\sqrt{5}} \langle 1, -1 \mid 1, -1 \rangle + \sqrt{\frac{3}{5}} \langle 1, -1 \mid 1, 0 \rangle + \frac{1}{\sqrt{5}} \langle 1, -1 \mid 1, 1 \rangle \right|^2$$
$$= \frac{1}{5}, \tag{5.285}$$

since $\langle 1, -1 \mid 1, 0 \rangle = \langle 1, -1 \mid 1, 1 \rangle = 0$ and $\langle 1, -1 \mid 1, -1 \rangle = 1$. Similarly, we can verify that the probabilities of measuring $l_z = 0$ and $\hbar$ are respectively given by

$$P_0 = \left| \langle 1, 0 \mid \psi \rangle \right|^2 = \left| \sqrt{\frac{3}{5}} \langle 1, 0 \mid 1, 0 \rangle \right|^2 = \frac{3}{5}, \tag{5.286}$$

$$P_1 = \left| \langle 1, 1 \mid \psi \rangle \right|^2 = \left| \sqrt{\frac{1}{5}} \langle 1, 1 \mid 1, 1 \rangle \right|^2 = \frac{1}{5}. \tag{5.287}$$

(c) After measuring $l_z = -\hbar$, the system will be in the eigenstate $\mid lm \rangle = \mid 1, -1 \rangle$, that is, $\psi(\theta, \varphi) = Y_{1,-1}(\theta, \varphi)$. We need first to calculate the expectation values of $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_x^2$, and $\hat{L}_y^2$ using $\mid 1, -1 \rangle$. Symmetry requires that $\langle 1, -1 \mid \hat{L}_x \mid 1, -1 \rangle = \langle 1, -1 \mid \hat{L}_y \mid 1, -1 \rangle = 0$. The expectation values of $\hat{L}_x^2$ and $\hat{L}_y^2$ are equal, as shown in (5.60); they are given by

$$\langle \hat{L}_x^2 \rangle = \langle \hat{L}_y^2 \rangle = \frac{1}{2} [\langle \hat{L}^2 \rangle - \langle \hat{L}_z^2 \rangle] = \frac{\hbar^2}{2} \left[l(l+1) - m^2 \right] = \frac{\hbar^2}{2}; \tag{5.288}$$

in this relation, we have used the fact that $l = 1$ and $m = -1$. Hence

$$\Delta L_x = \sqrt{\langle \hat{L}_x^2 \rangle} = \frac{\hbar}{\sqrt{2}} = \Delta L_y, \tag{5.289}$$

and the uncertainties product $\Delta L_x \Delta L_y$ is given by

$$\Delta L_x \Delta L_y = \sqrt{\langle \hat{L}_x^2 \rangle \langle \hat{L}_y^2 \rangle} = \frac{\hbar^2}{2}. \tag{5.290}$$

Problem 5.10
Find the angle between the angular momentum $l = 4$ and the z-axis for all possible orientations.

Solution
Since $m_l = 0, \pm 1, \pm 2, \ldots, \pm l$ and the angle between the orbital angular momentum l and the z-axis is $\cos\theta_{m_l} = m_l/\sqrt{l(l+1)}$ we have

$$\theta_{m_l} = \cos^{-1}\left[\frac{m_l}{\sqrt{l(l+1)}}\right] = \cos^{-1}\left[\frac{m_l}{2\sqrt{5}}\right]; \tag{5.291}$$

hence

$$\theta_0 = \cos^{-1}(0) = 90°, \tag{5.292}$$

$$\theta_1 = \cos^{-1}\left[\frac{1}{2\sqrt{5}}\right] = 77.08°, \qquad \theta_2 = \cos^{-1}\left[\frac{2}{2\sqrt{5}}\right] = 63.43°, \tag{5.293}$$

$$\theta_3 = \cos^{-1}\left[\frac{3}{2\sqrt{5}}\right] = 47.87°, \qquad \theta_4 = \cos^{-1}\left[\frac{4}{2\sqrt{5}}\right] = 26.57°. \tag{5.294}$$

The angles for the remaining quantum numbers $m_4 = -1, -2, -3, -4$ can be inferred at once from the relation

$$\theta_{-m_l} = 180° - \theta_{m_l}, \tag{5.295}$$

hence

$$\theta_{-1} = 180° - 77.08° = 102.92°, \qquad \theta_{-2} = 180° - 63.43° = 116.57°, \tag{5.296}$$

$$\theta_{-3} = 180° - 47.87° = 132.13°, \qquad \theta_{-4} = 180° - 26.57° = 153.43°. \tag{5.297}$$

Problem 5.11
Using $[\hat{X}, \hat{P}] = i\hbar$, calculate the various commutation relations between the following operators[5]

$$\hat{T}_1 = \frac{1}{4}(\hat{P}^2 - \hat{X}^2), \qquad \hat{T}_2 = \frac{1}{4}(\hat{X}\hat{P} + \hat{P}\hat{X}), \qquad \hat{T}_3 = \frac{1}{4}(\hat{P}^2 + \hat{X}^2).$$

[5]N. Zettili and F. Villars, *Nucl. Phys.*, **A469**, 77 (1987).

Solution

The operators $\hat{T}_1$, $\hat{T}_2$, and $\hat{T}_3$ can be viewed as describing some sort of *collective* vibrations; $\hat{T}_3$ has the structure of a harmonic oscillator Hamiltonian. The first commutator can be calculated as follows:

$$[\hat{T}_1, \hat{T}_2] = \frac{1}{4}[\hat{P}^2 - \hat{X}^2, \hat{T}_2] = \frac{1}{4}[\hat{P}^2, \hat{T}_2] - \frac{1}{4}[\hat{X}^2, \hat{T}_2], \tag{5.298}$$

where, using the commutation relation $[\hat{X}, \hat{P}] = i\hbar$, we have

$$\begin{aligned}
[\hat{P}^2, \hat{T}_2] &= \frac{1}{4}[\hat{P}^2, \hat{X}\hat{P}] + \frac{1}{4}[\hat{P}^2, \hat{P}\hat{X}] \\
&= \frac{1}{4}\hat{P}[\hat{P}, \hat{X}\hat{P}] + \frac{1}{4}[\hat{P}, \hat{X}\hat{P}]\hat{P} + \frac{1}{4}\hat{P}[\hat{P}, \hat{P}\hat{X}] + \frac{1}{4}[\hat{P}, \hat{P}\hat{X}]\hat{P} \\
&= \frac{1}{4}\hat{P}[\hat{P}, \hat{X}]\hat{P} + \frac{1}{4}[\hat{P}, \hat{X}]\hat{P}^2 + \frac{1}{4}\hat{P}^2[\hat{P}, \hat{X}] + \frac{1}{4}\hat{P}[\hat{P}, \hat{X}]\hat{P} \\
&= -\frac{i\hbar}{4}\hat{P}^2 - \frac{i\hbar}{4}\hat{P}^2 - \frac{i\hbar}{4}\hat{P}^2 - \frac{i\hbar}{4}\hat{P}^2 = -i\hbar\hat{P}^2, \qquad (5.299) \\
[\hat{X}^2, \hat{T}_2] &= \frac{1}{4}[\hat{X}^2, \hat{X}\hat{P}] + \frac{1}{4}[\hat{X}^2, \hat{P}\hat{X}] \\
&= \frac{1}{4}\hat{X}[\hat{X}, \hat{X}\hat{P}] + \frac{1}{4}[\hat{X}, \hat{X}\hat{P}]\hat{X} + \frac{1}{4}\hat{X}[\hat{X}, \hat{P}\hat{X}] + \frac{1}{4}[\hat{X}, \hat{P}\hat{X}]\hat{X} \\
&= \frac{1}{4}\hat{X}^2[\hat{X}, \hat{P}] + \frac{1}{4}\hat{X}[\hat{X}, \hat{P}]\hat{X} + \frac{1}{4}\hat{X}[\hat{X}, \hat{P}]\hat{X} + \frac{1}{4}[\hat{X}, \hat{P}]\hat{X}^2 \\
&= \frac{i\hbar}{4}\hat{X}^2 + \frac{i\hbar}{4}\hat{X}^2 + \frac{i\hbar}{4}\hat{X}^2 + \frac{i\hbar}{4}\hat{X}^2 = i\hbar\hat{X}^2; \qquad (5.300)
\end{aligned}$$

hence

$$[\hat{T}_1, \hat{T}_2] = \frac{1}{4}[\hat{P}^2 - \hat{X}^2, \hat{T}_2] = -\frac{1}{4}(i\hbar\hat{P}^2 + i\hbar\hat{X}^2) = -i\hbar\hat{T}_3. \tag{5.301}$$

The second commutator is calculated as follows:

$$[\hat{T}_2, \hat{T}_3] = \frac{1}{4}[\hat{T}_2, \hat{P}^2 + \hat{X}^2] = \frac{1}{4}[\hat{T}_2, \hat{P}^2] + \frac{1}{4}[\hat{T}_2, \hat{X}^2], \tag{5.302}$$

where $[\hat{T}_2, \hat{P}^2]$ and $[\hat{T}_2, \hat{X}^2]$ were calculated in (5.299) and (5.300):

$$[\hat{T}_2, \hat{P}^2] = i\hbar\hat{P}^2, \qquad [\hat{T}_2, \hat{X}^2] = -i\hbar\hat{X}^2. \tag{5.303}$$

Thus, we have

$$[\hat{T}_2, \hat{T}_3] = \frac{1}{4}(i\hbar\hat{P}^2 - i\hbar\hat{X}^2) = i\hbar\hat{T}_1. \tag{5.304}$$

The third commutator is

$$[\hat{T}_3, \hat{T}_1] = \frac{1}{4}[\hat{T}_3, \hat{P}^2 - \hat{X}^2] = \frac{1}{4}[\hat{T}_3, \hat{P}^2] - \frac{1}{4}[\hat{T}_3, \hat{X}^2], \tag{5.305}$$

where

$$\begin{aligned}
[\hat{T}_3, \hat{P}^2] &= \frac{1}{4}[\hat{P}^2, \hat{P}^2] + \frac{1}{4}[\hat{X}^2, \hat{P}^2] = \frac{1}{4}[\hat{X}^2, \hat{P}^2] = \frac{1}{4}\hat{X}[\hat{X}, \hat{P}^2] + \frac{1}{4}[\hat{X}, \hat{P}^2]\hat{X} \\
&= \frac{1}{4}\hat{X}\hat{P}[\hat{X}, \hat{P}] + \frac{1}{4}\hat{X}[\hat{X}, \hat{P}]\hat{P} + \frac{1}{4}\hat{P}[\hat{X}, \hat{P}]\hat{X} + \frac{1}{4}[\hat{X}, \hat{P}]\hat{P}\hat{X} \\
&= \frac{i\hbar}{4}(2\hat{X}\hat{P} + 2\hat{P}\hat{X}) = \frac{i\hbar}{2}(\hat{X}\hat{P} + \hat{P}\hat{X}), \qquad (5.306)
\end{aligned}$$

$$[\hat{T}_3, \hat{X}^2] = \frac{1}{4}[\hat{P}^2, \hat{X}^2] + \frac{1}{4}[\hat{X}^2, \hat{X}^2] = \frac{1}{4}[\hat{P}^2, \hat{X}^2] = -\frac{i\hbar}{2}(\hat{X}\hat{P} + \hat{P}\hat{X}); \tag{5.307}$$

hence

$$\begin{aligned}[\hat{T}_3, \hat{T}_1] &= \frac{1}{4}[\hat{T}_3, \hat{P}^2] - \frac{1}{4}[\hat{T}_3, \hat{X}^2] = \frac{i\hbar}{8}(\hat{X}\hat{P} + \hat{P}\hat{X}) + \frac{i\hbar}{8}(\hat{X}\hat{P} + \hat{P}\hat{X}) \\ &= \frac{i\hbar}{4}(\hat{X}\hat{P} + \hat{P}\hat{X}) = i\hbar\hat{T}_2. \end{aligned} \tag{5.308}$$

In sum, the commutation relations between $\hat{T}_1$, $\hat{T}_2$, and $\hat{T}_3$ are

$$[\hat{T}_1, \hat{T}_2] = -i\hbar\hat{T}_3, \qquad [\hat{T}_2, \hat{T}_3] = i\hbar\hat{T}_1, \qquad [\hat{T}_3, \hat{T}_1] = i\hbar\hat{T}_2. \tag{5.309}$$

These relations are similar to those of ordinary angular momentum, save for the minus sign in $[\hat{T}_1, \hat{T}_2] = -i\hbar\hat{T}_3$.

Problem 5.12
Consider a particle whose wave function is

$$\psi(x, y, z) = \frac{1}{4\sqrt{\pi}} \frac{2z^2 - x^2 - y^2}{r^2} + \sqrt{\frac{3}{\pi}} \frac{xz}{r^2}.$$

(a) Calculate $\hat{L}^2\psi(x, y, z)$ and $\hat{L}_z\psi(x, y, z)$. Find the total angular momentum of this particle.
(b) Calculate $\hat{L}_+\psi(x, y, z)$ and $\langle\psi \mid \hat{L}_+ \mid \psi\rangle$.
(c) If a measurement of the z-component of the orbital angular momentum is carried out, find the probabilities corresponding to finding the results 0, $\hbar$, and $-\hbar$.
(d) What is the probability of finding the particle at the position $\theta = \pi/3$ and $\varphi = \pi/2$ within $d\theta = 0.03$ rad and $d\varphi = 0.03$ rad?

Solution
(a) Since $Y_{20}(x, y, z) = \sqrt{5/16\pi}(3z^2 - r^2)/r^2$ and $Y_{2,\pm1}(x, y, z) = \mp\sqrt{15/8\pi}(x \pm iy)z/r^2$, we can write

$$\frac{2z^2 - x^2 - y^2}{r^2} = \frac{3z^2 - r^2}{r^2} = \sqrt{\frac{16\pi}{5}}\, Y_{20} \qquad \text{and} \qquad \frac{xz}{r^2} = \sqrt{\frac{2\pi}{15}}\,(Y_{2,-1} - Y_{21}); \tag{5.310}$$

hence

$$\psi(x, y, z) = \frac{1}{4\sqrt{\pi}}\sqrt{\frac{16\pi}{5}}Y_{20} + \sqrt{\frac{3}{\pi}}\sqrt{\frac{2\pi}{15}}\,(Y_{2,-1} - Y_{21}) = \frac{1}{\sqrt{5}}Y_{20} + \sqrt{\frac{2}{5}}\,(Y_{2,-1} - Y_{21}). \tag{5.311}$$

Having expressed ψ in terms of the spherical harmonics, we can now easily write

$$\hat{L}^2\psi(x, y, z) = \frac{1}{\sqrt{5}}\hat{L}^2Y_{20} + \sqrt{\frac{2}{5}}\hat{L}^2\,(Y_{2,-1} - Y_{21}) = 6\hbar^2\psi(x, y, z) \tag{5.312}$$

and

$$\hat{L}_z\psi(x, y, z) = \frac{1}{\sqrt{5}}\hat{L}_zY_{20} + \sqrt{\frac{2}{5}}\hat{L}_z\,(Y_{2,-1} - Y_{21}) = -\hbar\sqrt{\frac{2}{5}}\hat{L}_z\,(Y_{2,-1} + Y_{21}). \tag{5.313}$$

This shows that $\psi(x,y,z)$ is an eigenstate of $\hat{L}^2$ with eigenvalue $6\hbar^2$; $\psi(x,y,z)$ is, however, not an eigenstate of $\hat{L}_z$. Thus the total angular momentum of the particle is

$$\sqrt{\langle\psi \mid \hat{L}^2 \mid \psi\rangle} = \sqrt{6}\hbar. \tag{5.314}$$

(b) Using the relation $\hat{L}_+ Y_{lm} = \hbar\sqrt{l(l+1) - m(m+1)}\,Y_{l\,m+1}$, we have

$$\hat{L}_+\psi(x,y,z) = \frac{1}{\sqrt{5}}\hat{L}_+ Y_{20} + \sqrt{\frac{2}{5}}\hat{L}_+ (Y_{2,-1} - Y_{21}) = \hbar\sqrt{\frac{6}{5}}Y_{21} + \hbar\sqrt{\frac{2}{5}}\left(\sqrt{6}Y_{20} - 2Y_{22}\right); \tag{5.315}$$

hence

$$\begin{aligned}\langle\psi \mid \hat{L}_+ \mid \psi\rangle &= \left[\frac{1}{\sqrt{5}}\langle 2,\,0| + \sqrt{\frac{2}{5}}(\langle 2,\,-1| - \langle 2,\,1|)\right] \\ &\quad\times\left[\hbar\sqrt{\frac{6}{5}}Y_{21} + \hbar\sqrt{\frac{2}{5}}\left(\sqrt{6}Y_{20} - 2Y_{22}\right)\right] \\ &= 0.\end{aligned} \tag{5.316}$$

(c) Since $|\psi\rangle = (1/\sqrt{5})Y_{20} + \sqrt{2/5}(Y_{2,-1} - Y_{21})$, a calculation of $\langle\psi \mid \hat{L}_z \mid \psi\rangle$ yields

$$\langle\psi \mid \hat{L}_z \mid \psi\rangle = 0, \quad \text{with probability} \quad P_0 = \frac{1}{5}, \tag{5.317}$$

$$\langle\psi \mid \hat{L}_z \mid \psi\rangle = -\hbar, \quad \text{with probability} \quad P_{-1} = \frac{2}{5}, \tag{5.318}$$

$$\langle\psi \mid \hat{L}_z \mid \psi\rangle = \hbar, \quad \text{with probability} \quad P_1 = \frac{2}{5}. \tag{5.319}$$

(d) Since $\psi(x,y,z) = (1/4\sqrt{\pi})(2z^2 - x^2 - y^2)/r^2 + \sqrt{3/\pi}xz/r^2$ can be written in terms of the spherical coordinates as

$$\psi(\theta,\varphi) = \frac{1}{4\sqrt{\pi}}(3\cos\theta^2 - 1) + \sqrt{\frac{3}{\pi}}\sin\theta\cos\theta\cos\varphi, \tag{5.320}$$

the probability of finding the particle at the position θ and φ is

$$P(\theta,\varphi) = \left|\psi(\theta,\varphi)\right|^2 \sin\theta d\theta d\varphi = \left[\frac{1}{4\sqrt{\pi}}(3\cos^2\theta - 1) + \sqrt{\frac{3}{\pi}}\sin\theta\cos\theta\cos\varphi\right]^2 \sin\theta d\theta d\varphi; \tag{5.321}$$

hence

$$P\left(\frac{\pi}{3},\frac{\pi}{2}\right) = \left[\frac{1}{4\sqrt{\pi}}\left(3\cos^2\frac{\pi}{3} - 1\right) + 0\right]^2 (0.03)^2 \sin\frac{\pi}{3} = 9.7\times 10^{-7}. \tag{5.322}$$

Problem 5.13
Consider a particle of spin $s = 3/2$.

(a) Find the matrices representing the operators $\hat{S}_z$, $\hat{S}_x$, $\hat{S}_y$, $\hat{S}_x^2$, and $\hat{S}_y^2$ within the basis of $\hat{S}^2$ and $\hat{S}_z$.

(b) Find the energy levels of this particle when its Hamiltonian is given by

$$\hat{H} = \frac{\varepsilon_0}{\hbar^2}(\hat{S}_x^2 - \hat{S}_y^2) - \frac{\varepsilon_0}{\hbar}\hat{S}_z,$$

where ε_0 is a constant having the dimensions of energy. Are these levels degenerate?

(c) If the system was initially in an eigenstate $|\psi_0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, find the state of the system at time t.

Solution

(a) Following the same procedure that led to (5.73) and (5.75), we can verify that for $s = \frac{3}{2}$ we have

$$S_z = \frac{\hbar}{2}\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}, \tag{5.323}$$

$$\hat{S}_- = \hbar\begin{pmatrix} 0 & 0 & 0 & 0 \\ \sqrt{3} & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}, \qquad \hat{S}_+ = \hbar\begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{5.324}$$

which, when combined with $\hat{S}_x = (\hat{S}_+ + \hat{S}_-)/2$ and $\hat{S}_y = i(\hat{S}_- - \hat{S}_+)/2$, lead to

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}, \qquad \hat{S}_y = \frac{i\hbar}{2}\begin{pmatrix} 0 & -\sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & -2 & 0 \\ 0 & 2 & 0 & -\sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}. \tag{5.325}$$

Thus, we have

$$\hat{S}_x^2 = \frac{\hbar^2}{4}\begin{pmatrix} 3 & 0 & 2\sqrt{3} & 0 \\ 0 & 7 & 0 & 2\sqrt{3} \\ 2\sqrt{3} & 0 & 7 & 0 \\ 0 & 2\sqrt{3} & 0 & 3 \end{pmatrix}, \quad \hat{S}_y^2 = \frac{\hbar^2}{4}\begin{pmatrix} 3 & 0 & -2\sqrt{3} & 0 \\ 0 & 7 & 0 & -2\sqrt{3} \\ -2\sqrt{3} & 0 & 7 & 0 \\ 0 & -2\sqrt{3} & 0 & 3 \end{pmatrix}. \tag{5.326}$$

(b) The Hamiltonian is then given by

$$H = \frac{\varepsilon_0}{\hbar^2}(\hat{S}_x^2 - \hat{S}_y^2) - \frac{\varepsilon_0}{\hbar}\hat{S}_z = \frac{1}{2}\varepsilon_0\begin{pmatrix} -3 & 0 & 2\sqrt{3} & 0 \\ 0 & -1 & 0 & 2\sqrt{3} \\ 2\sqrt{3} & 0 & 1 & 0 \\ 0 & 2\sqrt{3} & 0 & 3 \end{pmatrix}. \tag{5.327}$$

The diagonalization of this Hamiltonian yields the following energy values:

$$E_1 = -\frac{5}{2}\varepsilon_0, \qquad E_2 = -\frac{3}{2}\varepsilon_0, \qquad E_3 = \frac{3}{2}\varepsilon_0, \qquad E_4 = \frac{5}{2}\varepsilon_0. \tag{5.328}$$

The corresponding normalized eigenvectors are given by

$$|1\rangle = \frac{1}{2}\begin{pmatrix} -\sqrt{3} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \; |2\rangle = \frac{1}{2}\begin{pmatrix} 0 \\ -\sqrt{3} \\ 0 \\ 1 \end{pmatrix}, \; |3\rangle = \frac{1}{\sqrt{12}}\begin{pmatrix} \sqrt{3} \\ 0 \\ 3 \\ 0 \end{pmatrix}, \; |4\rangle = \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ 0 \\ \sqrt{3} \end{pmatrix}. \tag{5.329}$$

None of the energy levels is degenerate.

(c) Since the initial state $|\psi_0\rangle$ can be written in terms of the eigenvectors (5.329) as follows:

$$|\psi_0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = -\frac{\sqrt{3}}{2}|1\rangle + \frac{1}{2}|3\rangle; \tag{5.330}$$

the eigenfunction at a later time t is given by

$$\begin{aligned} |\psi(t)\rangle &= -\frac{\sqrt{3}}{2}|1\rangle e^{-iE_1 t/\hbar} + \frac{1}{2}|3\rangle e^{-iE_3 t/\hbar} \\ &= -\frac{\sqrt{3}}{4}\begin{pmatrix} -\sqrt{3} \\ 0 \\ 1 \\ 0 \end{pmatrix}\exp\left[\frac{5i\varepsilon_0 t}{2\hbar}\right] + \frac{1}{2\sqrt{12}}\begin{pmatrix} \sqrt{3} \\ 0 \\ 3 \\ 0 \end{pmatrix}\exp\left[-\frac{3i\varepsilon_0 t}{2\hbar}\right]. \end{aligned} \tag{5.331}$$

5.9 Exercises

Exercise 5.1

(a) Show the following commutation relations:

$$[\hat{Y}, \hat{L}_y] = 0, \qquad [\hat{Y}, \hat{L}_z] = i\hbar\hat{X}, \qquad [\hat{Y}, \hat{L}_x] = -i\hbar\hat{Z},$$

$$[\hat{Z}, \hat{L}_z] = 0, \qquad [\hat{Z}, \hat{L}_x] = i\hbar\hat{Y}, \qquad [\hat{Z}, \hat{L}_y] = -i\hbar\hat{X}.$$

(b) Using a cyclic permutation of xyz, apply the results of (a) to infer expressions for $[\hat{X}, \hat{L}_x]$, $[\hat{X}, \hat{L}_y]$, and $[\hat{X}, \hat{L}_z]$.

(c) Use the results of (a) and (b) to calculate $[\hat{R}^2, \hat{L}_x]$, $[\hat{R}^2, \hat{L}_y]$, and $[\hat{R}^2, \hat{L}_x]$, where $\hat{R}^2 = \hat{X}^2 + \hat{Y}^2 + \hat{Z}^2$.

Exercise 5.2

(a) Show the following commutation relations:

$$[\hat{P}_y, \hat{L}_y] = 0, \qquad [\hat{P}_y, \hat{L}_z] = i\hbar\hat{P}_x, \qquad [\hat{P}_y, \hat{L}_x] = -i\hbar\hat{P}_z,$$

$$[\hat{P}_z, \hat{L}_z] = 0, \qquad [\hat{P}_z, \hat{L}_x] = i\hbar\hat{P}_y, \qquad [\hat{P}_z, \hat{L}_y] = -i\hbar\hat{P}_x.$$

(b) Use the results of (a) to infer by means of a cyclic permutation the expressions for $[\hat{P}_x, \hat{L}_x]$, $[\hat{P}_x, \hat{L}_y]$, and $[\hat{P}_x, \hat{L}_z]$.

(c) Use the results of (a) and (b) to calculate $[\hat{P}^2, \hat{L}_x]$, $[\hat{P}^2, \hat{L}_y]$, and $[\hat{P}^2, \hat{L}_z]$, where $\hat{P}^2 = \hat{P}_x^2 + \hat{P}_y^2 + \hat{P}_z^2$.

Exercise 5.3

If $\hat{L}_\pm$ and $\hat{R}_\pm$ are defined by $\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y$ and $\hat{R}_\pm = \hat{X} \pm i\hat{Y}$, prove the following commutators: (a) $[\hat{L}_\pm, \hat{R}_\pm] = \pm 2\hbar\hat{Z}$ and (b) $[\hat{L}_\pm, \hat{R}_\mp] = 0$.

Exercise 5.4

If $\hat{L}_\pm$ and $\hat{R}_\pm$ are defined by $\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y$ and $\hat{R}_\pm = \hat{X} \pm i\hat{Y}$, prove the following commutators: (a) $[\hat{L}_\pm, \hat{Z}] = \mp\hbar\hat{R}_\pm$, (b) $[\hat{L}_z, \hat{R}_\pm] = \pm\hbar\hat{R}_\pm$, and (c) $[\hat{L}_z, \hat{Z}] = 0$.

Exercise 5.5

Prove the following two relations: $\hat{\vec{R}} \cdot \hat{\vec{L}} = 0$ and $\hat{\vec{P}} \cdot \hat{\vec{L}} = 0$.

Exercise 5.6

The Hamiltonian due to the interaction of a particle of spin $\vec{S}$ with a magnetic field $\vec{B}$ is given by $\hat{H} = -\vec{S} \cdot \vec{B}$ where $\vec{S}$ is the spin. Calculate the commutator $[\vec{S}, \hat{H}]$.

Exercise 5.7

Prove the following relation:

$$[\hat{L}_z, \cos\varphi] = i\hbar \sin\varphi,$$

where φ is the azimuthal angle.

Exercise 5.8

Prove the following relation:

$$[\hat{L}_z, \sin(2\varphi)] = 2i\hbar\left(\sin^2\varphi - \cos^2\varphi\right),$$

where φ is the azimuthal angle. *Hint:* $[\hat{A}, \hat{B}\hat{C}] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C}$.

Exercise 5.9

Using the properties of $\hat{J}_+$ and $\hat{J}_-$, calculate $|j, \pm j\rangle$ and $|j, \pm m\rangle$ as functions of the action of $\hat{J}_\pm$ on the states $|j, \pm m\rangle$ and $|j, \pm j\rangle$, respectively.

Exercise 5.10

Consider the operator $\hat{A} = \frac{1}{2}(\hat{J}_x\hat{J}_y + \hat{J}_y\hat{J}_x)$.

(a) Calculate the expectation value of $\hat{A}$ and $\hat{A}^2$ with respect to the state $|\, j, m\rangle$.

(b) Use the result of (a) to find an expression for $\hat{A}^2$ in terms of: $\hat{\vec{J}}^4$, $\hat{\vec{J}}^2$, $\hat{J}_z^2$, $\hat{J}_+^4$, $\hat{J}_-^4$.

Exercise 5.11
Consider the wave function

$$\psi(\theta,\varphi) = 3\sin\theta\cos\theta e^{i\varphi} - 2(1-\cos^2\theta)e^{2i\varphi}.$$

(a) Write $\psi(\theta,\varphi)$ in terms of the spherical harmonics.
(b) Write the expression found in (a) in terms of the Cartesian coordinates.
(c) Is $\psi(\theta,\varphi)$ an eigenstate of $\hat{\vec{L}}^2$ or $\hat{L}_z$?
(d) Find the probability of measuring $2\hbar$ for the z-component of the orbital angular momentum.

Exercise 5.12
Show that $\hat{L}_z(\cos^2\varphi - \sin^2\varphi + 2i\sin\varphi\cos\varphi) = 2\hbar e^{2i\varphi}$, where φ is the azimuthal angle.

Exercise 5.13
Find the expressions for the spherical harmonics $Y_{30}(\theta,\varphi)$ and $Y_{3,\pm1}(\theta,\varphi)$,

$$Y_{30}(\theta,\varphi) = \sqrt{7/16\pi}(5\cos^3\theta - 3\cos\theta), \quad Y_{3,\pm1}(\theta,\varphi) = \mp\sqrt{21/64\pi}\sin\theta(5\cos^2\theta - 1)e^{\pm i\varphi},$$

in terms of the Cartesian coordinates x, y, z.

Exercise 5.14
(a) Show that the following expectation values between $|lm\rangle$ states satisfy the relations $\langle\hat{L}_x\rangle = \langle\hat{L}_y\rangle = 0$ and $\langle\hat{L}_x^2\rangle = \langle\hat{L}_y^2\rangle = \frac{1}{2}\left[l(l+1)\hbar^2 - m^2\hbar^2\right]$.
(b) Verify the inequality $\Delta L_x \Delta L_y \geq \hbar^2 m/2$, where $\Delta L_x = \sqrt{\langle L_x^2\rangle - \langle L_x\rangle^2}$.

Exercise 5.15
A particle of mass m is fixed at one end of a rigid rod of negligible mass and length R. The other end of the rod rotates in the xy plane about a bearing located at the origin, whose axis is in the z-direction.
(a) Write the system's total energy in terms of its angular momentum L.
(b) Write down the time-independent Schrödinger equation of the system. *Hint:* In spherical coordinates, only φ varies.
(c) Solve for the possible energy levels of the system, in terms of m and the moment of inertia $I = mR^2$.
(d) Explain why there is no zero-point energy (i.e., explain why the lowest energy is zero).

Exercise 5.16
Consider a system which is described by the state

$$\psi(\theta,\varphi) = \sqrt{\frac{3}{8}}Y_{11}(\theta,\varphi) + \sqrt{\frac{1}{8}}Y_{10}(\theta,\varphi) + AY_{1,-1}(\theta,\varphi),$$

where A is a real constant
(a) Calculate A so that $|\psi\rangle$ is normalized.
(b) Find $\hat{L}_+\psi(\theta,\varphi)$.
(c) Calculate the expectation values of $\hat{L}_x$ and $\hat{\vec{L}}^2$ in the state $|\psi\rangle$.

(d) Find the probability associated with a measurement that gives zero for the z-component of the angular momentum.

(e) Calculate $\langle\Phi|\hat{L}_z \mid \psi\rangle$ and $\langle\Phi|\hat{L}_- \mid \psi\rangle$ where

$$\Phi(\theta, \varphi) = \sqrt{\frac{8}{15}}Y_{11}(\theta, \varphi) + \sqrt{\frac{4}{15}}Y_{10}(\theta, \varphi) + \sqrt{\frac{3}{15}}Y_{2,-1}(\theta, \varphi).$$

Exercise 5.17

(a) Using the commutation relations of angular momentum, verify the validity of the (Jacobi) identity: $[\hat{J}_x, [\hat{J}_y, \hat{J}_z]] + [\hat{J}_y, [\hat{J}_z, \hat{J}_x]] + [\hat{J}_z, [\hat{J}_x, \hat{J}_y]] = 0$.

(b)Prove the following identity: $[\hat{J}_x^2, \hat{J}_y^2] = [\hat{J}_y^2, \hat{J}_z^2] = [\hat{J}_z^2, \hat{J}_x^2]$.

(c) Calculate the expressions of $\hat{L}_-\hat{L}_+Y_{lm}(\theta, \varphi)$ and $\hat{L}_+\hat{L}_-Y_{lm}(\theta, \varphi)$, and then infer the commutator $[\hat{L}_+\hat{L}_-, \hat{L}_-\hat{L}_+]Y_{lm}(\theta, \varphi)$.

Exercise 5.18

Consider a particle whose wave function is given by $\psi(x, y, z) = A[(x + z)y + z^2]/r^2 - A/3$, where A is a constant.

(a) Is ψ an eigenstate of $\hat{\vec{L}}^2$? If yes, what is the corresponding eigenvalue? Is it also an eigenstate of $\hat{L}_z$?

(b) Find the constant A so that ψ is normalized.

(c) Find the relative probabilities for measuring the various values of $\hat{L}_z$ and $\hat{\vec{L}}^2$, and then calculate the expectation values of $\hat{L}_z$ and $\hat{\vec{L}}^2$.

(d) Calculate $\hat{L}_\pm \mid \psi\rangle$ and then infer $\langle\psi \mid \hat{L}_\pm \mid \psi\rangle$.

Exercise 5.19

Consider a system which is in the state

$$\psi(\theta, \varphi) = \sqrt{\frac{2}{13}}Y_{3,-3} + \sqrt{\frac{3}{13}}Y_{3,-2} + \sqrt{\frac{3}{13}}Y_{30} + \sqrt{\frac{3}{13}}Y_{32} + \sqrt{\frac{2}{13}}Y_{33}.$$

(a) If $\hat{L}_z$ were measured, what values will one obtain and with what probabilities?

(b) If after a measurement of $\hat{L}_z$ we find $l_z = 2\hbar$, calculate the uncertainties ΔL_x and ΔL_y and their product $\Delta L_x \Delta L_y$.

(c) Find $\langle\psi \mid \hat{L}_x \mid \psi\rangle$ and $\langle\psi \mid \hat{L}_y \mid \psi\rangle$.

Exercise 5.20

(a) Calculate the energy eigenvalues of an axially symmetric rotor and find the degeneracy of each energy level (i.e., for each value of the azimuthal quantum number m, find how many states $|l\ m\rangle$ correspond to the same energy). We may recall that the Hamiltonian of an axially symmetric rotor is given by

$$\hat{H} = \frac{\hat{L}_x^2 + \hat{L}_y^2}{2I_1} + \frac{\hat{L}_z^2}{2I_2},$$

where I_1 and I_2 are the moments of inertia.

(b) From part (a) infer the energy eigenvalues for the various levels of $l = 3$.

(c) In the case of a rigid rotor (i.e., $I_1 = I_2 = I$), find the energy expression and the corresponding degeneracy relation.

(d) Calculate the orbital quantum number l and the corresponding energy degeneracy for a rigid rotor where the magnitude of the total angular momentum is $\sqrt{56}\hbar$.

Exercise 5.21
Consider a system of total angular momentum $j = 1$. We are interested here in the measurement of $\hat{J}_y$; its matrix is given by

$$\hat{J}_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}.$$

(a) What are the possible values will we obtain when measuring $\hat{J}_y$?
(b) Calculate $\langle \hat{J}_z \rangle$, $\langle \hat{J}_z^2 \rangle$, and ΔJ_z if the system is in the state $j_y = \hbar$.
(c) Repeat (b) for $\langle \hat{J}_x \rangle$, $\langle \hat{J}_x^2 \rangle$, and ΔJ_x.
(d) Verify that the uncertainties on $\hat{J}_x$ and $\hat{J}_z$ satisfy Heisenberg's principle; that is, show that

$$\Delta J_x \Delta J_z \geq \frac{1}{2} \left| \langle [\hat{J}_x, \hat{J}_z] \rangle \right|.$$

Exercise 5.22
Calculate $Y_{3,\pm 2}(\theta, \varphi)$ by applying the ladder operators $\hat{L}_\pm$ on $Y_{3,\pm 1}(\theta, \varphi)$.

Exercise 5.23
Consider a system of total angular momentum $j = 1$. We want to carry out measurements on

$$\hat{J}_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(a) What are the possible values will we obtain when measuring $\hat{J}_z$?
(b) Calculate $\langle \hat{J}_x \rangle$, $\langle \hat{J}_x^2 \rangle$, and ΔJ_x if the system is in the state $j_z = -\hbar$.
(c) Repeat (b) for $\langle \hat{J}_y \rangle$, $\langle \hat{J}_y^2 \rangle$, and ΔJ_y.

Exercise 5.24
Consider a system which is in the state

$$\psi(x, y, z) = \frac{1}{4\sqrt{\pi}} \frac{z}{r} + \frac{1}{\sqrt{3\pi}} \frac{x}{r}.$$

(a) Express $\psi(x, y, z)$ in terms of the spherical harmonics then calculate $\hat{L}^2 \psi(x, y, z)$ and $\hat{L}_z \psi(x, y, z)$. Is $\psi(x, y, z)$ an eigenstate of $\hat{L}^2$ or $\hat{L}_z$?
(b) Calculate $\hat{L}_\pm \psi(x, y, z)$ and $\langle \psi \mid \hat{L}_\pm \mid \psi \rangle$.
(c) If a measurement of the z-component of the orbital angular momentum is carried out, find the probabilities corresponding to finding the results 0, $\hbar$, and $-\hbar$.

Exercise 5.25
Consider a system whose wave function is given by

$$\psi(\theta, \varphi) = \frac{1}{2} Y_{00}(\theta, \varphi) + \frac{1}{\sqrt{3}} Y_{11}(\theta, \varphi) + \frac{1}{2} Y_{1,-1}(\theta, \varphi) + \frac{1}{\sqrt{6}} Y_{22}(\theta, \varphi).$$

(a) Is $\psi(\theta, \varphi)$ normalized?

(b) Is $\psi(\theta, \varphi)$ an eigenstate of $\hat{\vec{L}}^2$ or $\hat{L}_z$?

(c) Calculate $\hat{L}_\pm \psi(\theta, \varphi)$ and $\langle \psi \mid \hat{L}_\pm \mid \psi \rangle$.

(d) If a measurement of the z-component of the orbital angular momentum is carried out, find the probabilities corresponding to finding the results 0, $\hbar$, $-\hbar$, and $2\hbar$.

Exercise 5.26

Using the expression of $\hat{L}_-$ in spherical coordinates, prove the following two commutators: (a) $[\hat{L}_-, e^{-i\varphi} \sin\theta] = 0$ and (b) $[\hat{L}_-, \cos\theta] = \hbar e^{-i\varphi} \sin\theta$.

Exercise 5.27

Consider a particle whose angular momentum is $l = 1$.

(a) Find the eigenvalues and eigenvectors, $|1, m_x\rangle$, of $\hat{L}_x$.

(b) Express the state $|1, m_x = 1\rangle$ as a linear superposition of the eigenstates of $\hat{L}_z$. *Hint:* you need first to find the eigenstates of L_x and find which of them corresponds to the eigenvalue $m_x = 1$; this eigenvector will be expanded in the z basis.

(c) What is the probability of measuring $m_z = 1$ when the particle is in the eigenstate $|1, m_x = 1\rangle$? What about the probability corresponding to measuring $m_z = 0$?

(d) Suppose that a measurement of the z-component of angular momentum is performed and that the result $m_z = 1$ is obtained. Now we measure the x-component of angular momentum. What are the possible results and with what probabilities?

Exercise 5.28

Consider a system which is given in the following angular momentum eigenstates $| l, m\rangle$:

$$| \psi\rangle = \frac{1}{\sqrt{7}} | 1, -1\rangle + A | 1, 0\rangle + \sqrt{\frac{2}{7}} |1, 1\rangle,$$

where A is a real constant

(a) Calculate A so that $| \psi\rangle$ is normalized.

(b) Calculate the expectation values of $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$, and $\hat{\vec{L}}^2$ in the state $| \psi\rangle$.

(c) Find the probability associated with a measurement that gives $1\hbar$ for the z-component of the angular momentum.

(d) Calculate $\langle 1, m|\hat{L}_+^2 \mid \psi\rangle$ and $\langle 1, m|\hat{L}_-^2 \mid \psi\rangle$.

Exercise 5.29

Consider a particle of angular momentum $j = 3/2$.

(a) Find the matrices representing the operators $\hat{\vec{J}}^2$, $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$ in the $\{|\frac{3}{2}, m\rangle\}$ basis.

(b) Using these matrices, show that $\hat{J}_x$, $\hat{J}_y$, $\hat{J}_z$ satisfy the commutator $[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$.

(c) Calculate the mean values of $\hat{J}_x$ and $\hat{J}_x^2$ with respect to the state $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$.

(d) Calculate $\Delta J_x \Delta J_y$ with respect to the state

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

and verify that this product satisfies Heisenberg's uncertainty principle.

Exercise 5.30
Consider the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(a) Verify that $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$, where I is the unit matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(b) Calculate the commutators $[\sigma_x, \sigma_y]$, $[\sigma_x, \sigma_z]$, and $[\sigma_y, \sigma_z]$.
(c) Calculate the anticommutator $\sigma_x\sigma_y + \sigma_y\sigma_x$.
(d) Show that $e^{i\theta\sigma_y} = I\cos\theta + i\sigma_y \sin\theta$, where I is the unit matrix.
(e) Derive an expression for $e^{i\theta\sigma_z}$ by analogy with the one for σ_y.

Exercise 5.31
Consider a spin $\frac{3}{2}$ particle whose Hamiltonian is given by

$$\hat{H} = \frac{\varepsilon_0}{\hbar^2}(\hat{S}_x^2 - \hat{S}_y^2) - \frac{\varepsilon_0}{\hbar^2}\hat{S}_z^2,$$

where ε_0 is a constant having the dimensions of energy.
(a) Find the matrix of the Hamiltonian and diagonalize it to find the energy levels.
(b) Find the eigenvectors and verify that the energy levels are doubly degenerate.

Exercise 5.32
Find the energy levels of a spin $\frac{5}{2}$ particle whose Hamiltonian is given by

$$\hat{H} = \frac{\varepsilon_0}{\hbar^2}(\hat{S}_x^2 + \hat{S}_y^2) + \frac{\varepsilon_0}{\hbar}\hat{S}_z,$$

where ε_0 is a constant having the dimensions of energy. Are the energy levels degenerate?

Exercise 5.33
Consider an electron whose spin direction is located in the xy plane.
(a) Find the eigenvalues (call them λ_1, λ_2) and eigenstates ($|\lambda_1\rangle$, $|\lambda_2\rangle$) of the electron's spin operator $\vec{S}$.
(b) Assuming that the initial state of the electron is given by

$$|\psi_0\rangle = \frac{1}{3}|\lambda_1\rangle + \frac{2\sqrt{2}}{3}|\lambda_2\rangle,$$

find the probability of obtaining a value of $\hat{S} = -\hbar/2$ after measuring the spin of the electron.

Exercise 5.34
(a) Find the eigenvalues (call them λ_1, λ_2) and eigenstates ($|\lambda_1\rangle$, $|\lambda_2\rangle$) of the spin operator

$\vec{S}$ of an electron when $\vec{S}$ is pointing along an arbitrary unit vector $\vec{n}$ that lies within the yz plane.

(b) Assuming that the initial state of the electron is given by

$$|\psi_0\rangle = \frac{1}{2}|\lambda_1\rangle + \frac{\sqrt{3}}{2}|\lambda_2\rangle,$$

find the probability of obtaining a value of $\hat{S} = \hbar/2$ after measuring the spin of the electron.

Exercise 5.35
Consider a particle of spin $\frac{3}{2}$. Find the matrix for the component of the spin along a unit vector with arbitrary direction $\vec{n}$. Find its eigenvalues and eigenvectors. *Hint:*

$$\vec{n} = (\sin\theta\cos\varphi)\hat{\imath} + (\sin\theta\sin\varphi)\hat{\jmath} + (\cos\theta)\hat{k}.$$

Exercise 5.36
Show that $[\hat{J}_x\hat{J}_y, \hat{J}_z] + [\hat{J}_x, \hat{J}_y\hat{J}_z] = i\hbar\left(\hat{J}_x^2 - 2\hat{J}_y^2 + \hat{J}_z^2\right)$.

Exercise 5.37
Find the eigenvalues of the operators $\hat{\vec{L}}^2$ and $\hat{L}_z$ for each of the following states:
(a) $Y_{21}(\theta,\varphi)$,
(b) $Y_{3,-2}(\theta,\varphi)$,
(c) $\frac{1}{\sqrt{2}}[Y_{33}(\theta,\varphi) + Y_{3,-3}(\theta,\varphi)]$, and
(d) $Y_{40}(\theta,\varphi)$.

Exercise 5.38
Use the following general relations:

$$|\psi_x\rangle_\pm = \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2},\frac{1}{2}\right\rangle \pm \left|\frac{1}{2},-\frac{1}{2}\right\rangle\right], \qquad |\psi_y\rangle_\pm = \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2},\frac{1}{2}\right\rangle \pm i\left|\frac{1}{2},-\frac{1}{2}\right\rangle\right]$$

to verify the following eigenvalue equations:

$$\hat{S}_x|\psi_x\rangle_\pm = \pm\frac{\hbar}{2}|\psi_x\rangle_\pm \qquad \text{and} \qquad \hat{S}_y|\psi_y\rangle_\pm = \pm\frac{\hbar}{2}|\psi_y\rangle_\pm.$$

Chapter 6

Three-Dimensional Problems

6.1 Introduction

In this chapter we want to primarily examine how to solve the Schrödinger equation for spinless particles moving in three-dimensional potentials. We carry out this study in two different coordinate systems: the Cartesian system and the spherical system.

First, working within the context of Cartesian coordinates, we study the motion of a particle in different potentials: the free particle, a particle in a (three-dimensional) rectangular potential, and a particle in a harmonic oscillator potential. This study is going to be a simple generalization of the one-dimensional problems presented in Chapter 4. Unlike the one-dimensional case, three-dimensional problems often exhibit degeneracy, which occurs whenever the potential displays symmetry.

Second, using spherical coordinates, we describe the motion of a particle in spherically symmetric potentials. After presenting a general treatment, we consider several applications ranging from the free particle and the isotropic harmonic oscillator to the hydrogen atom.

Third, we show how to calculate the energy levels of a hydrogen atom when placed in a constant magnetic field; this gives rise to the Zeeman effect.

We conclude the chapter by looking at how the Schrödinger equation yields the the Pauli equation for a spin 1/2 particle in an electromagnetic field.

6.2 3D Problems in Cartesian Coordinates

We examine here how to extend Schrödinger's theory of one-dimensional problems (Chapter 4) to three dimensions.

6.2.1 General Treatment: Separation of Variables

The time-dependent Schrödinger equation for a spinless particle of mass m moving under the influence of a three-dimensional potential is

$$-\frac{\hbar^2}{2m}\vec{\nabla}^2\Psi(x,y,z,t)+\hat{V}(x,y,z,t)\Psi(x,y,z,t)=i\hbar\frac{\partial\Psi(x,y,z,t)}{\partial t}, \tag{6.1}$$

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

where $\vec{\nabla}^2$ is the Laplacian, $\vec{\nabla}^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$. As seen in Chapter 4, the wave function of a particle moving in a time-independent potential can be written as a product of spatial and time components:

$$\Psi(x, y, z, t) = \psi(x, y, z)e^{-iEt/\hbar}, \tag{6.2}$$

where $\psi(x, y, z)$ is the solution to the time-independent Schrödinger equation:

$$-\frac{\hbar^2}{2m}\vec{\nabla}^2\psi(x, y, z) + \hat{V}(x, y, z)\psi(x, y, z) = E\psi(x, y, z), \tag{6.3}$$

which is of the form $\hat{H}\psi = E\psi$.

This partial differential equation is generally difficult to solve. But, for those cases where the potential $\hat{V}(x, y, z)$ separates into the sum of three independent, one-dimensional terms (which should not be confused with a vector)

$$V(x, y, z) = V_x(x) + V_y(y) + V_z(z), \tag{6.4}$$

we can solve (6.3) by means of the technique of *separation of variables*. This technique consists of separating the three-dimensional Schrödinger equation (6.3) into three independent one-dimensional Schrödinger equations. Let us examine how to achieve this. Note that (6.3), in conjunction with (6.4), can be written as

$$\left[\hat{H}_x + \hat{H}_y + \hat{H}_z\right]\psi(x, y, z) = E\psi(x, y, z), \tag{6.5}$$

where $\hat{H}_x$ is given by

$$\hat{H}_x = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V_x(x); \tag{6.6}$$

the expressions for $\hat{H}_y$ and $\hat{H}_z$ are analogous.

As $\hat{V}(x, y, z)$ separates into three independent terms, we can also write $\psi(x, y, z)$ as a product of three functions of a single variable each:

$$\psi(x, y, z) = X(x)Y(y)Z(z). \tag{6.7}$$

Substituting (6.7) into (6.5) and dividing by $X(x)Y(y)Z(z)$, we obtain

$$\left[-\frac{\hbar^2}{2m}\frac{1}{X}\frac{d^2X}{dx^2} + V_x(x)\right] + \left[-\frac{\hbar^2}{2m}\frac{1}{Y}\frac{d^2Y}{dy^2} + V_y(y)\right] + \left[-\frac{\hbar^2}{2m}\frac{1}{Z}\frac{d^2Z}{dz^2} + V_z(z)\right] = E. \tag{6.8}$$

Since each expression in the square brackets depends on only one of the variables x, y, z, and since the sum of these three expressions is equal to a constant, E, each separate expression must then be equal to a constant such that the sum of these three constants is equal to E. For instance, the x-dependent expression is given by

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V_x(x)\right]X(x) = E_xX(x). \tag{6.9}$$

Similar equations hold for the y and z coordinates, with

$$E_x + E_y + E_z = E. \tag{6.10}$$

The separation of variables technique consists in essence of reducing the three-dimensional Schrödinger equation (6.3) into three separate one-dimensional equations (6.9).

6.2.2 The Free Particle

In the simple case of a free particle, the Schrödinger equation (6.3) reduces to three equations similar to (6.9) with $V_x = 0$, $V_y = 0$, and $V_z = 0$. The x-equation can be obtained from (6.9):

$$\frac{d^2X(x)}{dx^2} = -k_x^2 X(x), \tag{6.11}$$

where $k_x^2 = 2mE_x/\hbar^2$, and hence $E_x = \hbar^2 k_x^2/(2m)$. As shown in Chapter 4, the normalized solutions to (6.11) are plane waves

$$X(x) = \frac{1}{\sqrt{2\pi}} e^{ik_x x}. \tag{6.12}$$

Thus, the solution to the three-dimensional Schrödinger equation (6.3) is given by

$$\psi_{\vec{k}}(x, y, z) = (2\pi)^{-3/2} e^{ik_x x} e^{ik_y y} e^{ik_z z} = (2\pi)^{-3/2} e^{i\vec{k}\cdot\vec{r}}, \tag{6.13}$$

where $\vec{k}$ and $\vec{r}$ are the wave and position vectors of the particle, respectively. As for the total energy E, it is equal to the sum of the eigenvalues of the three one-dimensional equations (6.11):

$$E = E_x + E_y + E_z = \frac{\hbar^2}{2m}\left(k_x^2 + k_y^2 + k_z^2\right) = \frac{\hbar^2}{2m}\vec{k}^2. \tag{6.14}$$

Note that, since the energy (6.14) depends only on the magnitude of $\vec{k}$, all different orientations of $\vec{k}$ (obtained by varying k_x, k_y, k_z) subject to the condition

$$|\vec{k}| = \sqrt{k_x^2 + k_y^2 + k_z^2} = constant \tag{6.15}$$

generate different eigenfunctions (6.13) without a change in the energy. As the total number of orientations of $\vec{k}$ which preserve its magnitude is infinite, the energy of a free particle is *infinitely degenerate.*

Note that the solutions to the time-dependent Schrödinger equation (6.1) are obtained by substituting (6.13) into (6.2):

$$\Psi_{\vec{k}}(\vec{r}, t) = \psi(\vec{r}) e^{-i\omega t} = (2\pi)^{-3/2} e^{i(\vec{k}\cdot\vec{r} - \omega t)}, \tag{6.16}$$

where $\omega = E/\hbar$; this represents a propagating wave with wave vector $\vec{k}$. The orthonormality condition of this wave function is expressed by

$$\int \Psi^*_{\vec{k}'}(\vec{r}, t)\Psi_{\vec{k}}(\vec{r}, t)\, d^3r = \int \psi^*_{\vec{k}'}(\vec{r})\psi_{\vec{k}}(\vec{r})\, d^3r = (2\pi)^{-3} \int e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} d^3r = \delta(\vec{k} - \vec{k}'), \tag{6.17}$$

which can be written in Dirac's notation as

$$\langle \Psi_{\vec{k}'}(t)|\Psi_{\vec{k}}(t)\rangle = \langle \psi_{\vec{k}'}|\psi_{\vec{k}}\rangle = \delta(\vec{k}-\vec{k}'). \tag{6.18}$$

The free particle can be represented, as seen in Chapter 3, by a wave packet (a superposition of wave functions corresponding to the various wave vectors):

$$\Psi(\vec{r},t) = (2\pi)^{-3/2}\int A(\vec{k},t)\Psi_{\vec{k}}(\vec{r},t)\,d^3k = (2\pi)^{-3/2}\int A(\vec{k},t)e^{i(\vec{k}\cdot\vec{r}-\omega t)}d^3k, \tag{6.19}$$

where $A(\vec{k},t)$ is the Fourier transform of $\Psi(\vec{r},t)$:

$$A(\vec{k},t) = (2\pi)^{-3/2}\int \Psi(\vec{r},t)e^{-i(\vec{k}\cdot\vec{r}-\omega t)}d^3r. \tag{6.20}$$

As seen in Chapters 1 and 4, the position of the particle can be represented classically by the center of the wave packet.

6.2.3 The Box Potential

We are going to begin with the rectangular box potential, which has no symmetry, and then consider the cubic potential, which displays a great deal of symmetry, since the xyz axes are equivalent.

6.2.3.1 The Rectangular Box Potential

Consider first the case of a spinless particle of mass m confined in a *rectangular* box of sides a, b, c:

$$V(x,y,z) = \begin{cases} 0, & 0<x<a,\ 0<y<b,\ 0<z<c, \\ \infty, & \text{elsewhere,} \end{cases} \tag{6.21}$$

which can be written as $V(x,y,z) = V_x(x) + V_y(y) + V_z(z)$, with

$$V_x(x) = \begin{cases} 0, & 0<x<a, \\ \infty, & \text{elsewhere;} \end{cases} \tag{6.22}$$

the potentials $V_y(y)$ and $V_z(z)$ have similar forms.

The wave function $\psi(x,y,z)$ must vanish at the walls of the box. We have seen in Chapter 4 that the solutions for this potential are of the form

$$X(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n_x\pi}{a}x\right), \qquad n_x = 1,2,3,\ldots, \tag{6.23}$$

and the corresponding energy eigenvalues are

$$E_{n_x} = \frac{\hbar^2\pi^2}{2ma^2}n_x^2. \tag{6.24}$$

From these expressions we can write the normalized three-dimensional eigenfunctions and their corresponding energies:

$$\boxed{\psi_{n_x n_y n_z}(x,y,z) = \sqrt{\frac{8}{abc}}\sin\left(\frac{n_x\pi}{a}x\right)\sin\left(\frac{n_y\pi}{b}y\right)\sin\left(\frac{n_z\pi}{c}z\right),} \tag{6.25}$$

$$\boxed{E_{n_x n_y n_z} = \frac{\hbar^2\pi^2}{2m}\left(\frac{n_x^2}{a^2}+\frac{n_y^2}{b^2}+\frac{n_z^2}{c^2}\right).} \tag{6.26}$$

6.2.3.2 The Cubic Potential

For the simpler case of a *cubic* box of side L, the energy expression can be inferred from (6.26) by substituting $a = b = c = L$:

$$\boxed{E_{n_x n_y n_z} = \frac{\hbar^2\pi^2}{2mL^2}(n_x^2+n_y^2+n_z^2), \qquad n_x,\ n_y,\ n_z = 1,2,3,\ldots.} \tag{6.27}$$

The ground state corresponds to $n_x = n_y = n_z = 1$; its energy is given by

$$E_{111} = \frac{3\pi^2\hbar^2}{2mL^2} = 3E_1, \tag{6.28}$$

where, as shown in Chapter 4, $E_1 = \pi^2\hbar^2/(2mL^2)$ is the zero-point energy of a particle in a one-dimensional box. Thus, the zero-point energy for a particle in a three-dimensional box is three times that in a one-dimensional box. The factor 3 can be viewed as originating from the fact that we are confining the particle symmetrically in all three dimensions.

The first excited state has three possible sets of quantum numbers $(n_x,\ n_y,\ n_z) = (2,1,1)$, $(1,2,1)$, $(1,1,2)$ corresponding to three different states $\psi_{211}(x,y,z)$, $\psi_{121}(x,y,z)$, and $\psi_{112}(x,y,z)$, where

$$\psi_{211}(x,y,z) = \sqrt{\frac{8}{L^3}}\sin\left(\frac{2\pi}{L}x\right)\sin\left(\frac{\pi}{L}y\right)\sin\left(\frac{\pi}{L}z\right); \tag{6.29}$$

the expressions for $\psi_{121}(x,y,z)$ and $\psi_{112}(x,y,z)$ can be inferred from $\psi_{211}(x,y,z)$. Notice that all three states have the same energy:

$$E_{211} = E_{121} = E_{112} = 6\frac{\pi^2\hbar^2}{2mL^2} = 6E_1. \tag{6.30}$$

The first excited state is thus threefold *degenerate*.

Degeneracy occurs only when there is a symmetry in the problem. For the present case of a particle in a cubic box, there is a great deal of symmetry, since all three dimensions are equivalent. Note that for the rectangular box where none of the sides a, b, and c is equal to the other two, there is no degeneracy since the three dimensions are not equivalent. Moreover, degeneracy did not exist when we treated one-dimensional problems in Chapter 4, for they give rise to only one quantum number.

The second excited state also has three different states, and hence it is threefold degenerate; its energy is equal to $9E_1$: $E_{221} = E_{212} = E_{122} = 9E_1$.

The energy spectrum is shown in Table 6.1, where every nth level is characterized by its energy, its quantum numbers, and its degeneracy g_n.

Table 6.1: Energy levels and their degeneracies for the cubic potential, with $E_1 = \frac{\pi^2\hbar^2}{2mL^2}$.

$E_{n_x n_y n_z}/E_1$	(n_x, n_y, n_z)	g_n
3	(111)	1
6	(211), (121), (112)	3
9	(221), (212), (122)	3
11	(311), (131), (113)	3
12	(222)	1
14	(321), (312), (231), (213), (132), (123)	6

6.2.4 The Harmonic Oscillator

We are going to begin with the anisotropic oscillator, which displays no symmetry, and then consider the isotropic oscillator where the xyz axes are all equivalent.

6.2.4.1 The Anisotropic Oscillator

Consider a particle of mass m moving in a three-dimensional anisotropic oscillator potential

$$\hat{V}(\hat{x}, \hat{y}, \hat{z}) = \frac{1}{2}m\omega_x^2\hat{x}^2 + \frac{1}{2}m\omega_y^2\hat{y}^2 + \frac{1}{2}m\omega_z^2\hat{z}^2. \tag{6.31}$$

Its Schrödinger equation separates into three equations similar to (6.9):

$$-\frac{\hbar^2}{2m}\frac{d^2X(x)}{dx^2} + \frac{1}{2}m\omega_x^2x^2X(x) = E_xX(x), \tag{6.32}$$

with similar equations for $Y(y)$ and $Z(z)$. The eigenenergies corresponding to the potential (6.31) can be expressed as

$$\boxed{E_{n_x n_y n_z} = E_{n_x} + E_{n_y} + E_{n_z} = \left(n_x + \frac{1}{2}\right)\hbar\omega_x + \left(n_y + \frac{1}{2}\right)\hbar\omega_y + \left(n_z + \frac{1}{2}\right)\hbar\omega_z,} \tag{6.33}$$

with $n_x, n_y, n_z = 0, 1, 2, 3, \ldots$. The corresponding stationary states are

$$\psi_{n_x n_y n_z}(x, y, z) = X_{n_x}(x)Y_{n_y}(y)Z_{n_z}(z), \tag{6.34}$$

where $X_{n_x}(x)$, $Y_{n_y}(y)$, and $Z_{n_z}(z)$ are one-dimensional harmonic oscillator wave functions. These states are not degenerate, because the potential (6.31) has no symmetry (it is anisotropic).

6.2.4.2 The Isotropic Harmonic Oscillator

Consider now an *isotropic* harmonic oscillator potential. Its energy eigenvalues can be inferred from (6.33) by substituting $\omega_x = \omega_y = \omega_z = \omega$,

$$\boxed{E_{n_x n_y n_z} = \left(n_x + n_y + n_z + \frac{3}{2}\right)\hbar\omega.} \tag{6.35}$$

Table 6.2: Energy levels and their degeneracies for an isotropic harmonic oscillator.

n	$2E_n/(\hbar\omega)$	$(n_x n_y n_z)$	g_n
0	3	(000)	1
1	5	(100), (010), (001)	3
2	7	(200), (020), (002)	6
		(110), (101), (011)	
3	9	(300), (030), (003)	10
		(210), (201), (021)	
		(120), (102), (012)	
		(111)	

Since the energy depends on the sum of n_x, n_y, n_z, any set of quantum numbers having the same sum will represent states of equal energy.

The ground state, whose energy is $E_{000} = 3\hbar\omega/2$, is not degenerate. The first excited state is threefold degenerate, since there are three different states, ψ_{100}, ψ_{010}, ψ_{001}, that correspond to the same energy $5\hbar\omega/2$. The second excited state is sixfold degenerate; its energy is $7\hbar\omega/2$.

In general, we can show that the degeneracy g_n of the nth excited state, which is equal to the number of ways the nonnegative integers n_x, n_y, n_z may be chosen to total to n, is given by

$$\boxed{g_n = \frac{1}{2}(n+1)(n+2),} \tag{6.36}$$

where $n = n_x + n_y + n_z$. Table 6.2 displays the first few energy levels along with their degeneracies.

Example 6.1 (Degeneracy of a harmonic oscillator)
Show how to derive the degeneracy relation (6.36).

Solution
For a fixed value of n, the degeneracy g_n is given by the number of ways of choosing n_x, n_y, and n_z so that $n = n_x + n_y + n_z$.

For a fixed value of n_x, the number of ways of choosing n_y and n_z so that $n_y + n_z = n - n_x$ is given by $(n - n_x + 1)$; this can be shown as follows. For a given value of n_x, the various permissible values of (n_y, n_z) are given by $(n_y, n_z) = (0, n - n_x)$, $(1, n - n_x - 1)$, $(2, n - n_x - 2)$, $(3, n - n_x - 3)$, ..., $(n - n_x - 3, 3)$, $(n - n_x - 2, 2)$, $(n - n_x - 1, 1)$, and $(n - n_x, 0)$. In all, there are $(n - n_x + 1)$ sets of (n_y, n_z) so that $n_y + n_z = n - n_x$. Now, since the values of n_x can vary from 0 to n, the degeneracy is then given by

$$g_n = \sum_{n_x=0}^{n}(n - n_x + 1) = (n+1)\sum_{n_x=0}^{n} 1 - \sum_{n_x=0}^{n} n_x = (n+1)^2 - \frac{1}{2}n(n+1) = \frac{1}{2}(n+1)(n+2). \tag{6.37}$$

A more primitive way of calculating this series is to use Gauss's method: simply write the series $\sum_{n_x=0}^{n}(n - n_x + 1)$ in the following two equivalent forms:

$$g_n = (n+1) + n + (n-1) + (n-2) + \cdots + 4 + 3 + 2 + 1, \tag{6.38}$$

$$g_n = 1 + 2 + 3 + 4 + \cdots + (n-2) + (n-1) + n + (n+1). \tag{6.39}$$

Since both of these two series contain $(n + 1)$ terms each, a term by term addition of these relations yields

$$\begin{aligned} 2g_n &= (n+2) + (n+2) + (n+2) + \cdots + (n+2) + (n+2) + (n+2) \\ &= (n+1)(n+2); \end{aligned} \tag{6.40}$$

hence $g_n = \frac{1}{2}(n + 1)(n + 2)$.

6.3 3D Problems in Spherical Coordinates

6.3.1 Central Potential: General Treatment

In this section we study the structure of the Schrödinger equation for a particle of mass[1] M moving in a *spherically symmetric* potential

$$V(\vec{r}) = V(r), \tag{6.41}$$

which is also known as the *central* potential.

The time-independent Schrödinger equation for this particle, of momentum $-i\hbar\vec{\nabla}$ and position vector $\vec{r}$, is

$$\left[-\frac{\hbar^2}{2M}\nabla^2 + V(r)\right]\psi(\vec{r}) = E\psi(\vec{r}). \tag{6.42}$$

Since the Hamiltonian is spherically symmetric, we are going to use the spherical coordinates (r, θ, φ) which are related to their Cartesian counterparts by

$$x = r\sin\theta\cos\varphi, \qquad y = r\sin\theta\sin\varphi, \qquad z = r\cos\theta. \tag{6.43}$$

The Laplacian ∇^2 separates into a radial part ∇_r^2 and an angular part ∇_Ω^2 as follows (see Chapter 5):

$$\nabla^2 = \nabla_r^2 - \frac{1}{\hbar^2 r^2}\nabla_\Omega^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{\hbar^2 r^2}\hat{\vec{L}}^2 = \frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{1}{\hbar^2 r^2}\hat{\vec{L}}^2, \tag{6.44}$$

where $\hat{\vec{L}}$ is the orbital angular momentum with

$$\hat{\vec{L}}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]. \tag{6.45}$$

[1] Throughout this section we will designate the mass of the particle by a capital M to avoid any confusion with the azimuthal quantum number m.

In spherical coordinates the Schrödinger equation therefore takes the form

$$\left[-\frac{\hbar^2}{2M}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{2Mr^2}\hat{\vec{L}}^2 + V(r)\right]\psi(\vec{r}) = E\psi(\vec{r}). \tag{6.46}$$

The first term of this equation can be viewed as the *radial* kinetic energy

$$-\frac{\hbar^2}{2M}\frac{1}{r}\frac{\partial^2}{\partial r^2}r = \frac{\hat{p}_r^2}{2M}, \tag{6.47}$$

since the radial momentum operator is given by the Hermitian form[2]

$$\hat{p}_r = \frac{1}{2}\left[\left(\frac{\vec{r}}{r}\right)\cdot\hat{\vec{p}} + \hat{\vec{p}}\cdot\left(\frac{\vec{r}}{r}\right)\right] = -i\hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right) \equiv -i\hbar\frac{1}{r}\frac{\partial}{\partial r}r. \tag{6.48}$$

The second term $\hat{\vec{L}}^2/(2Mr^2)$ of (6.46) can be identified with the *rotational* kinetic energy, for this term is generated from a "pure" rotation of the particle about the origin (i.e., no change in the radial variable r, where Mr^2 is its moment of inertia with respect to the origin).

Now, since $\hat{\vec{L}}^2$ as shown in (6.45) does not depend on r, it commutes with both $\hat{V}(r)$ and the radial kinetic energy; hence it also commutes with the Hamiltonian $\hat{H}$. In addition, since $\hat{L}_z$ commutes with $\hat{\vec{L}}^2$, the three operators $\hat{H}$, $\hat{\vec{L}}^2$, and L_z mutually commute:

$$[\hat{H}, \hat{\vec{L}}^2] = [\hat{H}, \hat{L}_z] = 0. \tag{6.49}$$

Thus $\hat{H}$, $\hat{\vec{L}}^2$, and $\hat{L}_z$ have common eigenfunctions. We have seen in Chapter 5 that the simultaneous eigenfunctions of $\hat{\vec{L}}^2$ and $\hat{L}_z$ are given by the spherical harmonics $Y_{lm}(\theta,\varphi)$:

$$\hat{\vec{L}}^2 Y_{lm}(\theta,\varphi) = l(l+1)\hbar^2 Y_{lm}(\theta,\varphi), \tag{6.50}$$

$$\hat{L}_z Y_{lm}(\theta,\varphi) = m\hbar Y_{lm}(\theta,\varphi). \tag{6.51}$$

Since the Hamiltonian in (6.46) is a sum of a radial part and an angular part, we can look for solutions that are products of a radial part and an angular part, where the angular part is simply the spherical harmonic $Y_{lm}(\theta,\varphi)$:

$$\psi(\vec{r}) = \langle\vec{r} \mid nlm\rangle = \psi_{nlm}(r,\theta,\varphi) = R_{nl}(r)Y_{lm}(\theta,\varphi). \tag{6.52}$$

Note that the orbital angular momentum of a system moving in a central potential is conserved, since, as shown in (6.49), it commutes with the Hamiltonian.

The radial wave function $R_{nl}(r)$ has yet to be found. The quantum number n is introduced to identify the eigenvalues of $\hat{H}$:

$$\hat{H} \mid nlm\rangle = E_n \mid nlm\rangle. \tag{6.53}$$

[2]Note that we can show that the commutator between the position operator, $\hat{r}$, and the radial momentum operator, $\hat{p}_r$, is given by: $[\hat{r}, \hat{p}_r] = i\hbar$ (the proof is left as an exercise).

Substituting (6.52) into (6.46) and using the fact that $\psi_{nlm}(r,\theta,\varphi)$ is an eigenfunction of $\hat{\vec{L}}^2$ with eigenvalue $l(l+1)\hbar^2$, then dividing through by $R_{nl}(r)Y_{lm}(\theta,\varphi)$ and multiplying by $2Mr^2$, we end up with an equation where the radial and angular degrees of freedom are separated:

$$\left[-\hbar^2\frac{r}{R_{nl}}\frac{\partial^2}{\partial r^2}(rR_{nl})+2Mr^2(V(r)-E)\right]+\left[\frac{\hat{\vec{L}}^2Y_{lm}(\theta,\varphi)}{Y_{lm}(\theta,\varphi)}\right]=0. \tag{6.54}$$

The terms inside the first square bracket are independent of θ and φ and those of the second are independent of r. They must then be separately equal to constants and their sum equal to zero. The second square bracket is nothing but (6.50), the eigenvalue equation of $\hat{\vec{L}}^2$; hence it is equal to $l(l+1)\hbar^2$. As for the first bracket, it must be equal to $-l(l+1)\hbar^2$; this leads to an equation known as the *radial equation* for a central potential:

$$\boxed{-\frac{\hbar^2}{2M}\frac{d^2}{dr^2}(rR_{nl}(r))+\left[V(r)+\frac{l(l+1)\hbar^2}{2Mr^2}\right](rR_{nl}(r))=E_n(rR_{nl}(r)).} \tag{6.55}$$

Note that (6.55), which gives the energy levels of the system, does not depend on the azimuthal quantum number m. Thus, the energy E_n is $(2l+1)$-fold *degenerate*. This is due to the fact that, for a given l, there are $(2l+1)$ different eigenfunctions ψ_{nlm} (i.e., $\psi_{nl\,-l}$, $\psi_{nl\,-l+1}$, $\ldots$, $\psi_{nl\,l-1}$, $\psi_{nl\,l}$) which correspond to the same eigenenergy E_n. This degeneracy property is peculiar to central potentials.

Note that (6.55) has the structure of a one-dimensional equation in r,

$$-\frac{\hbar^2}{2M}\frac{d^2U_{nl}(r)}{dr^2}+\left[V(r)+\frac{l(l+1)\hbar^2}{2Mr^2}\right]U_{nl}(r)=E_nU_{nl}(r), \tag{6.56}$$

or

$$\boxed{-\frac{\hbar^2}{2M}\frac{d^2U_{nl}(r)}{dr^2}+V_{\text{eff}}(r)U_{nl}(r)=E_nU_{nl}(r),} \tag{6.57}$$

whose solutions give the energy levels of the system. The wave function $U_{nl}(r)$ is given by

$$U_{nl}(r)=rR_{nl}(r) \tag{6.58}$$

and the potential by

$$V_{\text{eff}}(r)=V(r)+\frac{l(l+1)\hbar^2}{2Mr^2}, \tag{6.59}$$

which is known as the *effective* or *centrifugal* potential, where $V(r)$ is the central potential and $l(l+1)\hbar^2/2Mr^2$ is a repulsive or centrifugal potential, associated with the orbital angular momentum, which tends to repel the particle away from the center. As will be seen later, in the case of atoms, $V(r)$ is the Coulomb potential resulting from the attractive forces between the electrons and the nucleus. Notice that although (6.57) has the structure of a one-dimensional eigenvalue equation, it differs from the one-dimensional Schrödinger equation in one major aspect: the variable r cannot have negative values, for it varies from $r=0$ to $r\to+\infty$. We must therefore require the wave function $\psi_{nlm}(r,\theta,\varphi)$ to be finite for all values of r between 0 and ∞, notably for $r=0$. But if $R_{nl}(0)$ is finite, $rR_{nl}(r)$ must vanish at $r=0$, i.e.,

$$\lim_{r\to0}[rR_{nl}(r)]=U_{nl}(0)=0. \tag{6.60}$$

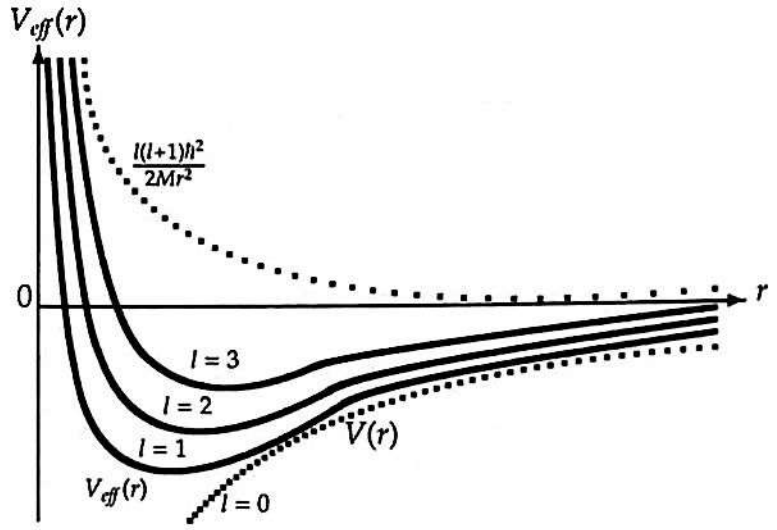


Figure 6.1: The effective potential $V_{eff}(r) = V(r) + \hbar^2 l(l+1)/(2Mr^2)$ corresponding to several values of l: $l = 0,1,2,3$; $V(r)$ is an attractive central potential, while $\hbar^2 l(l+1)/(2Mr^2)$ is a repulsive (centrifugal) potential.

Thus, to make the radial equation (6.57) equivalent to a one-dimensional eigenvalue problem, we need to assume that the particle's potential is given by the effective potential $V_{eff}(r)$ for $r > 0$ and by an infinite potential for $r \leq 0$.

For the eigenvalue equation (6.57) to describe bound states, the potential $V(r)$ must be attractive (i.e., negative) because $l(l+1)\hbar^2/(2Mr^2)$ is repulsive. Figure 6.1 shows that, as l increases, the depth of $V_{eff}(r)$ decreases and its minimum moves farther away from the origin. The farther the particle from the origin, the less bound it will be. This is due to the fact that as the particle's angular momentum increases, the particle becomes less and less bound.

In summary, we want to emphasize the fact that, for spherically symmetric potentials, the Schrödinger equation (6.46) reduces to a trivial angular equation (6.50) for $\hat{\vec{L}}^2$ and to a one-dimensional radial equation (6.57).

Remark
When a particle has orbital and spin degrees of freedom, its total wave function $|\Psi\rangle$ consists of a product of two parts: a spatial part, $\psi(\vec{r})$, and a spin part, $|s, m_s\rangle$; that is, $|\Psi\rangle = |\psi\rangle\,|s, m_s\rangle$. In the case of an electron moving in a central field, besides the quantum numbers n, l, m_l, a complete description of its state would require a fourth quantum number, the spin quantum number m_s: $|nlm_lm_s\rangle = |nlm_l\rangle\,|s, m_s\rangle$; hence

$$\Psi_{nlm_lm_s}(\vec{r}) = \psi_{nlm_l}(\vec{r})\,|s, m_s\rangle = R_{nl}(r)Y_{lm_l}(\theta,\varphi)\,|s, m_s\rangle. \tag{6.61}$$

Since the spin does not depend on the spatial degrees of freedom, the spin operator does not act on the spatial wave function $\psi_{nlm_l}(\vec{r})$ but acts only on the spin part $|s, m_s\rangle$; conversely, $\hat{\vec{L}}$ acts only the spatial part.

Table 6.3: First few spherical Bessel and Neumann functions.

Bessel functions $j_l(r)$	Neumann functions $n_l(r)$
$j_0(r) = \frac{\sin r}{r}$	$n_0(r) = -\frac{\cos r}{r}$
$j_1(r) = \frac{\sin r}{r^2} - \frac{\cos r}{r}$	$n_1(r) = -\frac{\cos r}{r^2} - \frac{\sin r}{r}$
$j_2(r) = \left(\frac{3}{r^3} - \frac{1}{r}\right)\sin r - \frac{3\cos r}{r^2}$	$n_2(r) = -\left(\frac{3}{r^3} - \frac{1}{r}\right)\cos r - \frac{3}{r^2}\sin r$

6.3.2 The Free Particle in Spherical Coordinates

In what follows we want to apply the general formalism developed above to study the motion of a free particle of mass M and energy $E_k = \hbar^2k^2/(2M)$, where k is the wave number ($k = |\vec{k}|$). The Hamiltonian $\hat{H} = -\hbar^2\nabla^2/(2M)$ of a free particle commutes with $\hat{\vec{L}}^2$ and $\hat{L}_z$. Since $V(r) = 0$ the Hamiltonian of a free particle is *rotationally invariant*. The free particle can then be viewed as a special case of central potentials. We have shown above that the radial and angular parts of the wave function can be separated, $\psi_{klm}(r,\theta,\varphi) = \langle r\theta\varphi \mid klm\rangle = R_{kl}(r)Y_{lm}(\theta,\varphi)$.

The radial equation for a free particle is obtained by setting $V(r) = 0$ in (6.55):

$$-\frac{\hbar^2}{2M}\frac{1}{r}\frac{d^2}{dr^2}(rR_{kl}(r)) + \frac{l(l+1)\hbar^2}{2Mr^2}R_{kl}(r) = E_kR_{kl}(r), \tag{6.62}$$

which can be rewritten as

$$-\frac{1}{r}\frac{d^2}{dr^2}(rR_{kl}(r)) + \frac{l(l+1)}{r^2}R_{kl}(r) = k^2R_{kl}(r), \tag{6.63}$$

where $k^2 = 2ME_k/\hbar^2$.

Using the change of variable $\rho = kr$, we can reduce this equation to

$$\frac{d^2\mathcal{R}_l(\rho)}{d\rho^2} + \frac{2}{\rho}\frac{d\mathcal{R}_l(\rho)}{d\rho} + \left[1 - \frac{l(l+1)}{\rho^2}\right]\mathcal{R}_l(\rho) = 0, \tag{6.64}$$

where $\mathcal{R}_l(\rho) = \mathcal{R}_l(kr) = R_{kl}(r)$. This differential equation is known as the *spherical Bessel equation*. The general solutions to this equation are given by an independent linear combination of the *spherical Bessel functions* $j_l(\rho)$ and the *spherical Neumann functions* $n_l(\rho)$:

$$\mathcal{R}_l(\rho) = A_lj_l(\rho) + B_ln_l(\rho), \tag{6.65}$$

where $j_l(\rho)$ and $n_l(\rho)$ are given by

$$j_l(\rho) = (-\rho)^l\left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l\frac{\sin\rho}{\rho}, \qquad n_l(\rho) = -(-\rho)^l\left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l\frac{\cos\rho}{\rho}. \tag{6.66}$$

The first few spherical Bessel and Neumann functions are listed in Table 6.3 and their shapes are displayed in Figure 6.2.

Expanding $\sin\rho/\rho$ and $\cos\rho/\rho$ in a power series of ρ, we see that the functions $j_l(\rho)$ and $n_l(\rho)$ reduce for small values of ρ (i.e., near the origin) to

$$j_l(\rho) \simeq \frac{2^ll!}{(2l+1)!}\rho^l, \qquad n_l(\rho) \simeq -\frac{(2l)!}{2^ll!}\rho^{-l-1}, \qquad \rho \ll 1, \tag{6.67}$$

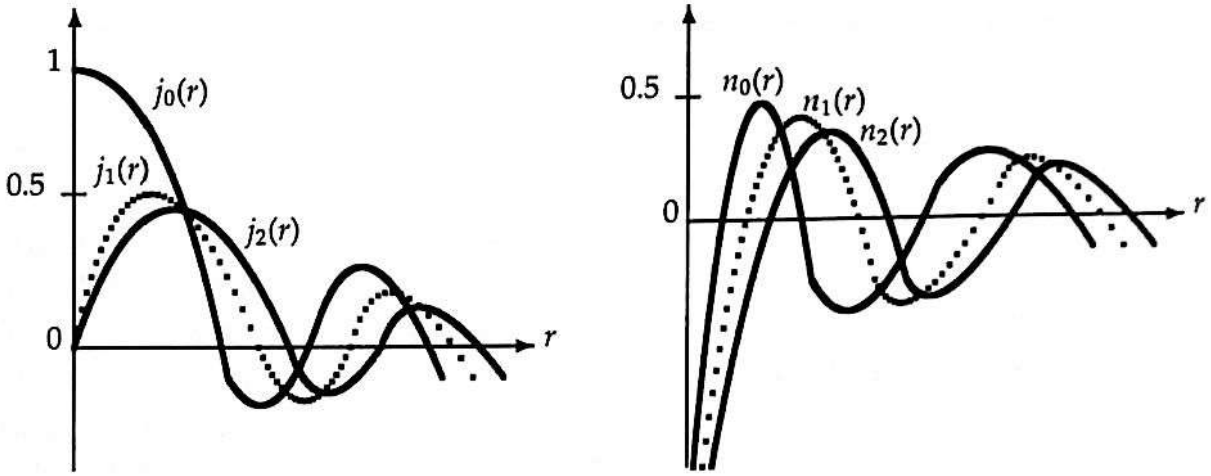


Figure 6.2: Spherical Bessel functions $j_l(r)$ and spherical Neumann functions $n_l(r)$; only the Bessel functions are finite at the origin.

and for large values of ρ to

$$j_l(\rho) \simeq \frac{1}{\rho}\sin\left(\rho - \frac{l\pi}{2}\right), \qquad n_l(\rho) \simeq -\frac{1}{\rho}\cos\left(\rho - \frac{l\pi}{2}\right), \qquad \rho \gg 1. \tag{6.68}$$

Since the Neumann functions $n_l(\rho)$ diverge at the origin, and since the wave functions ψ_{klm} are required to be finite everywhere in space, the functions $n_l(\rho)$ are unacceptable solutions to the problem. Hence only the spherical Bessel functions $j_l(kr)$ contribute to the eigenfunctions of the free particle:

$$\psi_{klm}(r,\theta,\psi) = j_l(kr)Y_{lm}(\theta,\varphi), \tag{6.69}$$

where $k = \sqrt{2ME_k}/\hbar$. As shown in Figure 6.2, the amplitude of the wave functions becomes smaller and smaller as r increases. At large distances, the wave functions are represented by spherical waves.

Note that, since the index k in $E_k = \hbar^2k^2/(2M)$ varies *continuously*, the energy spectrum of a free particle is *infinitely degenerate*. This is because all orientations of $\vec{k}$ in space correspond to the same energy.

Remark

We have studied the free particle within the context of Cartesian and spherical coordinate systems. Whereas the energy is given in both coordinate systems by the same expression, $E_k = \hbar^2k^2/(2M)$, the wave functions are given in Cartesian coordinates by *plane waves* $e^{i\vec{k}\cdot\vec{r}}$ (see (6.13)) and in spherical coordinates by *spherical waves* $j_l(kr)Y_{lm}(\theta,\varphi)$ (see (6.69)). We can, however, show that both sets of wave functions are equivalent, since we can express a plane wave $e^{i\vec{k}\cdot\vec{r}}$ in terms of spherical wave states $j_l(kr)Y_{lm}(\theta,\varphi)$. In particular, we can generate plane waves from a linear combination of spherical states that have the same k but different values of l and m:

$$e^{i\vec{k}\cdot\vec{r}} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_{lm}\, j_l(kr)Y_{lm}(\theta,\varphi). \tag{6.70}$$

The problem therefore reduces to finding the expansion coefficients a_{lm}. For instance, in the case where $\vec{k}$ is along the z-axis, $m = 0$, we can show that

$$e^{i\vec{k}\cdot\vec{r}} = e^{ikr\cos\theta} = \sum_{l=0}^{\infty} i^l\,(2l+1)j_l\,(kr)P_l(\cos\theta), \tag{6.71}$$

where $P_l(\cos\theta)$ are the Legendre polynomials, with $Y_{l0}(\theta,\varphi) \sim P_l(\cos\theta)$. The wave functions $\psi_{klm}(r,\theta,\varphi) = j_l(kr)Y_{lm}(\theta,\varphi)$ describe a free particle of energy E_k, with angular momentum l, but they give no information on the linear momentum $\vec{p}$ (ψ_{klm} is an eigenstate of $\hat{H}$, $\hat{\vec{L}}^2$, and $\hat{L}_z$ but not of $\hat{\vec{P}}$). On the other hand, the plane wave $e^{i\vec{k}\cdot\vec{r}}$ which is an eigenfunction of $\hat{H}$ and $\hat{\vec{P}}$, but not of $\hat{\vec{L}}^2$ nor of $\hat{L}_z$, gives no information about the particle's angular momentum. That is, plane waves describe states with well-defined linear momenta but poorly defined angular momenta. Conversely, spherical waves describe states with well-defined angular momenta but poorly defined linear momenta.

6.3.3 The Spherical Square Well Potential

Consider now the problem of a particle of mass M in an attractive square well potential

$$V(r) = \begin{cases} -V_0, & r < a, \\ 0, & r > a. \end{cases} \tag{6.72}$$

Let us consider the cases $0 < r < a$ and $r > a$ separately.

6.3.3.1 Case where $0 < r < a$

Inside the well, $0 < r < a$, the time-independent Schrödinger equation for this particle can be obtained from (6.55):

$$-\frac{\hbar^2}{2M}\frac{1}{r}\frac{d^2}{dr^2}\,(rR_l(r)) + \frac{l(l+1)\hbar^2}{2Mr^2}R_l(r) = (E+V_0)R_l(r). \tag{6.73}$$

Using the change of variable $\rho = k_1 r$, where k_1 is now given by $k_1 = \sqrt{2M(E+V_0)}/\hbar$, we see that (6.73) reduces to the spherical Bessel differential equation (6.64). As in the case of a free particle, the radial wave function must be finite everywhere, and is given as follows in terms of the spherical Bessel functions $j_l(k_1 r)$:

$$R_l(r) = Aj_l(k_1 r) = Aj_l\left(\frac{\sqrt{2M(E+V_0)}}{\hbar}\,r\right), \qquad \text{for} \qquad r < a, \tag{6.74}$$

where A is a normalization constant.

6.3.3.2 Case where $r > a$

Outside the well, $r > a$, the particle moves freely; its Schrödinger equation is (6.62):

$$-\frac{\hbar^2}{2M}\frac{1}{r}\frac{d^2}{dr^2}\,(rR_{kl}(r)) + \frac{l(l+1)\hbar^2}{2Mr^2}R_{kl}(r) = E_k R_{kl}(r) \qquad (r > a). \tag{6.75}$$

Two possibilities arise here, depending on whether the energy is negative or positive.

- The negative energy case corresponds to bound states (i.e., to a discrete energy spectrum). The general solutions of (6.75) are similar to those of (6.63), but k is now an imaginary number; that is, we must replace k by ik_2 and, hence, the solutions are given by linear combinations of $j_l(ik_2r)$ and $n_l(ik_2r)$:

$$R_l(ik_2r) = B\left[j_l(ik_2r) \pm n_l(ik_2r)\right], \tag{6.76}$$

where B is a normalization constant, with $k_2 = \sqrt{-2ME}/\hbar$. **Note:** Linear combinations of $j_l(\rho)$ and $n_l(\rho)$ can be expressed in terms of the *spherical Hankel functions* of the first kind, $h_l^{(1)}(\rho)$, and the second kind, $h_l^{(2)}(\rho)$, as follows:

$$h_l^{(1)}(\rho) = j_l(\rho) + in_l(\rho), \tag{6.77}$$

$$h_l^{(2)}(\rho) = j_l(\rho) - in_l(\rho) = \left(h_l^{(1)}(\rho)\right)^*. \tag{6.78}$$

The first few spherical Hankel functions of the first kind are

$$h_0^{(1)}(\rho) = -i\frac{e^{i\rho}}{\rho}, \quad h_1^{(1)}(\rho) = -\left(\frac{1}{\rho} + \frac{i}{\rho^2}\right)e^{i\rho}, \quad h_2^{(1)}(\rho) = \left(\frac{i}{\rho} - \frac{3}{\rho^2} - \frac{3i}{\rho^3}\right)e^{i\rho}. \tag{6.79}$$

The asymptotic behavior of the Hankel functions when $\rho \to \infty$ can be inferred from (6.68):

$$h_l^{(1)}(\rho) \to -\frac{i}{\rho}e^{i(\rho - l\pi/2)}, \qquad h_l^{(2)}(\rho) \to \frac{i}{\rho}e^{-i(\rho - l\pi/2)}. \tag{6.80}$$

The solutions that need to be retained in (6.76) must be finite everywhere. As can be inferred from Eq (6.80), only the Hankel functions of the first kind $h_l^{(1)}(ik_2r)$ are finite at large values of r (the functions $h_l^{(2)}(ik_2r)$ diverge for large values of r). Thus, the wave functions outside the well that are physically meaningful are those expressed in terms of the Hankel functions of the first kind (see (6.76)):

$$R_l(ik_2r) = Bh_l^{(1)}\left(i\frac{\sqrt{-2ME}}{\hbar}r\right) = Bj_l\left(i\frac{\sqrt{-2ME}}{\hbar}r\right) + iBn_l\left(i\frac{\sqrt{-2ME}}{\hbar}r\right). \tag{6.81}$$

The continuity of the radial function and its derivative at $r = a$ yields

$$\frac{1}{h_l^{(1)}(ik_2r)}\frac{dh_l^{(1)}(ik_2r)}{dr}\bigg|_{r=a} = \frac{1}{j_l(k_1r)}\frac{dj_l(k_1r)}{dr}\bigg|_{r=a}. \tag{6.82}$$

For the $l = 0$ states, this equation reduces to

$$-k_2 = k_1\cot(k_1a). \tag{6.83}$$

This continuity condition is analogous to the transcendental equation we obtained in Chapter 4 when we studied the one-dimensional finite square well potential.

- The positive energy case corresponds to the continuous spectrum (unbound or scattering states), where the solution is asymptotically oscillatory. The solution consists of a linear combination of $j_l(k'r)$ and $n_l(k'r)$, where $k' = \sqrt{2ME}/\hbar$. Since the solution must be finite everywhere, the continuity condition at $r = a$ determines the coefficients of the linear combination. The particle can move freely to infinity with a finite kinetic energy $E = \hbar^2k'^2/(2M)$.

6.3.4 The Isotropic Harmonic Oscillator

The radial Schrödinger equation for a particle of mass M in an isotropic harmonic oscillator potential

$$V(r) = \frac{1}{2}M\omega^2 r^2 \tag{6.84}$$

is obtained from (6.57):

$$-\frac{\hbar^2}{2M}\frac{d^2U_{nl}(r)}{dr^2} + \left[\frac{1}{2}M\omega^2 r^2 + \frac{l(l+1)\hbar^2}{2Mr^2}\right]U_{nl}(r) = EU_{nl}(r). \tag{6.85}$$

We are going to solve this equation by examining the behavior of the solutions at the asymptotic limits (at very small and very large values of r). On the one hand, when $r \to 0$, the E and $M\omega^2r^2/2$ terms become too small compared to the $l(l+1)\hbar^2/2Mr^2$ term. Hence, when $r \to 0$, Eq. (6.85) reduces to

$$-\frac{\hbar^2}{2M}\frac{d^2U(r)}{dr^2} + \frac{l(l+1)\hbar^2}{2Mr^2}U(r) = 0; \tag{6.86}$$

the solutions of this equation are of the form $U(r) \sim r^{l+1}$. On the other hand,when $r \to \infty$, the E and $l(l+1)\hbar^2/2Mr^2$ terms become too small compared to the $M\omega^2r^2/2$ term; hence, the asymptotic form of (6.85) when $r \to \infty$ is

$$-\frac{\hbar^2}{2M}\frac{d^2U(r)}{dr^2} + \frac{1}{2}M\omega^2 r^2 U(r) = 0, \tag{6.87}$$

which admits solutions of type $U(r) \sim e^{-M\omega r^2/2\hbar}$. Combining (6.86) and (6.87), we can write the solutions of (6.85) as

$$U(r) = f(r)r^{l+1}e^{-M\omega r^2/2\hbar}, \tag{6.88}$$

where $f(r)$ is a function of r. Substituting this expression into (6.85), we obtain an equation for $f(r)$:

$$\frac{d^2f(r)}{dr^2} + 2\left(\frac{l+1}{r} - \frac{M\omega}{\hbar}r\right)\frac{df(r)}{dr} + \left[\frac{2ME}{\hbar^2} - (2l+3)\frac{M\omega}{\hbar}\right]f(r) = 0. \tag{6.89}$$

Let us try a power series solution

$$f(r) = \sum_{n=0}^{\infty} a_n r^n = a_0 + a_1 r + a_2 r^2 + \cdots + a_n r^n + \cdots. \tag{6.90}$$

Substituting this function into (6.89), we obtain

$$\sum_{n=0}^{\infty}\Big\{n(n-1)a_n r^{n-2} + 2\left(\frac{l+1}{r} - \frac{M\omega}{\hbar}r\right)na_n r^{n-1} + \left[\frac{2ME}{\hbar^2} - (2l+3)\frac{M\omega}{\hbar}\right]a_n r^n\Big\} = 0, \tag{6.91}$$

which in turn reduces to

$$\sum_{n=0}^{\infty}\left\{n(n+2l+1)a_n r^{n-2} + \left[-\frac{2M\omega}{\hbar}n + \frac{2ME}{\hbar^2} - (2l+3)\frac{M\omega}{\hbar}\right]a_n r^n\right\} = 0. \tag{6.92}$$

For this equation to hold, the coefficients of the various powers of r must vanish separately. For instance, when $n = 0$ the coefficient of r^{-2} is indeed zero:

$$0 \cdot (2l + 1)a_0 = 0. \tag{6.93}$$

Note that a_0 need not be zero for this equation to hold. The coefficient of r^{-1} corresponds to $n = 1$ in (6.92); for this coefficient to vanish, we must have

$$1 \cdot (2l + 2)a_1 = 0. \tag{6.94}$$

Since $(2l + 2)$ cannot be zero, because the quantum number l is a positive integer, a_1 must vanish.

The coefficient of r^n results from the relation

$$\sum_{n=0} \left\{ (n + 2)(n + 2l + 3)a_{n+2} + \left[\frac{2ME}{\hbar^2} - \frac{M\omega}{\hbar}(2n + 2l + 3) \right] a_n \right\} r^n = 0, \tag{6.95}$$

which leads to the recurrence formula

$$(n + 2)(n + 2l + 3)a_{n+2} = \left[\frac{-2ME}{\hbar^2} + \frac{M\omega}{\hbar}(2n + 2l + 3) \right] a_n. \tag{6.96}$$

This recurrence formula shows that all coefficients a_n corresponding to odd values of n are zero, since $a_1 = 0$ (see (6.94)). The function $f(r)$ must therefore contain only even powers of r:

$$f(r) = \sum_{n=0}^{\infty} a_{2n} r^{2n} = \sum_{n'=0,2,4,\cdots}^{\infty} a_{n'} r^{n'}, \tag{6.97}$$

where all coefficients a_{2n}, with $n \geq 1$, are proportional to a_0.

Now note that when $n \to +\infty$ the function $f(r)$ diverges, for it behaves asymptotically like e^{r^2}. To obtain a finite solution, we must require the series (6.97) to stop at a maximum power $r^{n'}$; hence it must be *polynomial*. For this, we require $a_{n'+2}$ to be zero. Thus, setting $a_{n'+2} = 0$ into the recurrence formula (6.96) and since $a_{n'} \neq 0$, we obtain at once the *quantization condition*

$$2\frac{M}{\hbar^2} E_{n'l} - \frac{M\omega}{\hbar}(2n' + 2l + 3) = 0, \tag{6.98}$$

or

$$E_{n'l} = \left(n' + l + \frac{3}{2} \right) \hbar\omega, \tag{6.99}$$

where n' is even (see (6.97)). Denoting n' by $2N$, where $N = 0, 1, 2, 3, \ldots$, we rewrite this energy expression as

$$\boxed{E_n = \left(n + \frac{3}{2} \right) \hbar\omega \qquad (n = 0, 1, 2, 3, \ldots),} \tag{6.100}$$

where $n = n' + l = 2N + l$.

The ground state, whose energy is $E_0 = \frac{3}{2}\hbar\omega$, is not degenerate; the first excited state, $E_1 = \frac{5}{2}\hbar\omega$, is threefold degenerate; and the second excited state, $E_2 = \frac{7}{2}\hbar\omega$, is sixfold degenerate

Table 6.4: Energy levels E_n and degeneracies g_n for an isotropic harmonic oscillator.

n	E_n	Nl	m	g_n
0	$\frac{3}{2}\hbar\omega$	0 0	0	1
1	$\frac{5}{2}\hbar\omega$	0 1	$\pm 1, 0$	3
2	$\frac{7}{2}\hbar\omega$	1 0	0	6
		0 2	$\pm 2, \pm 1, 0$	
3	$\frac{9}{2}\hbar\omega$	1 1	$\pm 1, 0$	10
		0 3	$\pm 3, \pm 2, \pm 1, 0$	

(Table 6.4). As shown in the following example, the degeneracy relation for the nth level is given by

$$\boxed{g_n = \frac{1}{2}(n+1)(n+2).} \tag{6.101}$$

This expression is in agreement with (6.36) obtained for an isotropic harmonic oscillator in Cartesian coordinates.

Finally, since the radial wave function is given by $R_{nl}(r) = U_{nl}(r)/r$, where $U_{nl}(r)$ is listed in (6.88) with $f(r)$ being a polynomial in r^{2l} of degree $(n-l)/2$, the total wave function for the isotropic harmonic oscillator is

$$\psi_{nlm}(r,\theta,\varphi) = R_{nl}(r)Y_{lm}(\theta,\varphi) = \frac{U_{nl}(r)}{r}Y_{lm}(\theta,\varphi) = r^l f(r)Y_{lm}(\theta,\varphi)e^{-M\omega r^2/2\hbar}, \tag{6.102}$$

where l takes only odd or only even values. For instance, the ground state corresponds to $(n,l,m) = (0,0,0)$; its wave function is

$$\psi_{000}(r,\theta,\varphi) = R_{00}(r)Y_{00}(\theta,\varphi) = \frac{2}{\sqrt{\sqrt{\pi}}}\left(\frac{M\omega}{\hbar}\right)^{3/4} e^{-M\omega r^2/2\hbar}Y_{00}(\theta,\varphi). \tag{6.103}$$

The (n,l,m) configurations of the first, second, and third excited states can be determined as follows. The first excited state has three degenerate states: $(1,1,m)$ with $m = -1,0,1$. The second excited states has 6 degenerate states: $(2,0,0)$ and $(2,2,m)$ with $m = -2,-1,0,1,2$. The third excited state has 10 degenerate states: $(3,1,m)$ with $m = -1,0,1$ and $(3,3,m)$ where $m = -3,-2,-1,0,1,2,3$. Some of these wave functions are given by

$$\psi_{11m}(r,\theta,\varphi) = R_{11}(r)Y_{1m}(\theta,\varphi) = \sqrt{\frac{8}{3\sqrt{\pi}}}\left(\frac{M\omega}{\hbar}\right)^{5/4} re^{-M\omega r^2/2\hbar}Y_{1m}(\theta,\varphi), \tag{6.104}$$

$$\psi_{200}(r,\theta,\varphi) = R_{20}(r)Y_{00}(\theta,\varphi) = \sqrt{\frac{8}{3\sqrt{\pi}}}\left(\frac{M\omega}{\hbar}\right)^{3/4}\left(\frac{3}{2} - \frac{M\omega}{\hbar}r^2\right)e^{-M\omega r^2/2\hbar}Y_{00}(\theta,\varphi), \tag{6.105}$$

$$\psi_{31m}(r,\theta,\varphi) = R_{31}(r)Y_{1m}(\theta,\varphi) = \frac{4}{\sqrt{15\sqrt{\pi}}}\left(\frac{M\omega}{\hbar}\right)^{7/4} r^2 e^{-M\omega r^2/2\hbar}Y_{1m}(\theta,\varphi). \tag{6.106}$$

Example 6.2 (Degeneracy relation for an isotropic oscillator)
Prove the degeneracy relation (6.101) for an isotropic harmonic oscillator.

Solution
Since $n = 2N + l$ the quantum numbers n and l must have the same parity. Also, since the isotropic harmonic oscillator is spherically symmetric, its states have definite parity[3]. In addition, since the parity of the states corresponding to a central potential is given by $(-1)^l$, the quantum number l (hence n) can take only even or only odd values. Let us consider separately the cases when n is even or odd.

First, when n is even the degeneracy g_n of the nth excited state is given by

$$g_n = \sum_{l=0,2,4,\ldots}^{n}(2l+1) = \sum_{l=0,2,4,\ldots}^{n}1 + 2\sum_{l=0,2,4,\ldots}^{n} l = \frac{1}{2}(n+2) + \frac{n(n+2)}{2} = \frac{1}{2}(n+1)(n+2). \tag{6.107}$$

A more explicit way of obtaining this series consists of writing it in the following two equivalent forms:

$$g_n = 1 + 5 + 9 + 13 + \cdots + (2n-7) + (2n-3) + (2n+1), \tag{6.108}$$

$$g_n = (2n+1) + (2n-3) + (2n-7) + (2n-11) + \cdots + 13 + 9 + 5 + 1. \tag{6.109}$$

We then add them, term by term, to get

$$2g_n = (2n+2) + (2n+2) + (2n+2) + (2n+2) + \cdots + (2n+2) = (2n+2)\left(\frac{n}{2}+1\right). \tag{6.110}$$

This relation yields $g_n = \frac{1}{2}(n+1)(n+2)$, which proves (6.101) when n is even.

Second, when n is odd, a similar treatment leads to

$$g_n = \sum_{l=1,3,5,7,\ldots}^{n}(2l+1) = \sum_{l=1,3,5,7,\ldots}^{n}1 + 2\sum_{l=1,3,5,7,\ldots}^{n} l = \frac{1}{2}(n+1) + \frac{1}{2}(n+1)^2 = \frac{1}{2}(n+1)(n+2), \tag{6.111}$$

which proves (6.101) when n is odd. Note that this degeneracy relation is, as expected, identical with the degeneracy expression (6.36) obtained for a harmonic oscillator in Cartesian coordinates.

6.3.5 The Hydrogen Atom

The hydrogen atom consists of an electron and a proton. For simplicity, we will ignore their spins. The wave function then depends on six coordinates $\vec{r}_e(x_e, y_e, z_e)$ and $\vec{r}_p(x_p, y_p, z_p)$, where $\vec{r}_e$ and $\vec{r}_p$ are the electron and proton position vectors, respectively. According to the probabilistic interpretation of the wave function, the quantity $|\Psi(\vec{r}_e, \vec{r}_p, t)|^2 d^3r_e\, d^3r_p$ represents the probability that a simultaneous measurement of the electron and proton positions at time t will result in the electron being in the volume element d^3r_e and the proton in d^3r_p.

[3]Recall from Chapter 4 that if the potential of a system is symmetric, $V(x) = V(-x)$, the states of the system must be either odd or even.

The time-dependent Schrödinger equation for the hydrogen atom is given by

$$\left[-\frac{\hbar^2}{2m_p}\nabla_p^2 - \frac{\hbar^2}{2m_e}\nabla_e^2 + V(r)\right]\Psi(\vec{r}_e, \vec{r}_p, t) = i\hbar\frac{\partial}{\partial t}\Psi(\vec{r}_e, \vec{r}_p, t), \tag{6.112}$$

where ∇_p^2 and ∇_e^2 are the Laplacians with respect to the proton and the electron degrees of freedom, with $\nabla_p^2 = \partial^2/\partial x_p^2 + \partial^2/\partial y_p^2 + \partial^2/\partial z_p^2$ and $\nabla_e^2 = \partial^2/\partial x_e^2 + \partial^2/\partial y_e^2 + \partial^2/\partial z_e^2$, and where $V(r)$ is the potential (interaction) between the electron and the proton. This interaction, which depends only on the distance that separates the electron and the proton $\vec{r} = \vec{r}_e - \vec{r}_p$, is given by the Coulomb potential:

$$V(r) = -\frac{e^2}{r}. \tag{6.113}$$

Note: Throughout this text, we will be using the CGS units for the Coulomb potential where it is given by $V(r) = -e^2/r$ (in the MKS units, however, it is given by $V(r) = -e^2/(4\pi\varepsilon_0 r)$).

Since V does not depend on time, the solutions of (6.112) are stationary; hence, they can be written as follows:

$$\Psi(\vec{r}_e, \vec{r}_p, t) = \chi(\vec{r}_e, \vec{r}_p)e^{-iEt/\hbar}, \tag{6.114}$$

where E is the total energy of the electron–proton system. Substituting this into (6.112), we obtain the time-independent Schrödinger equation for the hydrogen atom:

$$\left[-\frac{\hbar^2}{2m_p}\nabla_p^2 - \frac{\hbar^2}{2m_e}\nabla_e^2 - \frac{e^2}{|\vec{r}_e - \vec{r}_p|}\right]\chi(\vec{r}_e, \vec{r}_p) = E\chi(\vec{r}_e, \vec{r}_p). \tag{6.115}$$

6.3.5.1 Separation of the Center of Mass Motion

Since V depends only on the relative distance r between the electron and proton, instead of the coordinates $\vec{r}_e$ and $\vec{r}_p$ (position vectors of the electron and proton), it is more appropriate to use the coordinates of the center of mass, $\vec{R} = X\hat{\imath} + Y\hat{\jmath} + Z\hat{k}$, and the relative coordinates of the electron with respect to the proton, $\vec{r} = x\hat{\imath} + y\hat{\jmath} + z\hat{k}$. The transformation from $\vec{r}_e, \vec{r}_p$ to $\vec{R}, \vec{r}$ is given by

$$\vec{R} = \frac{m_e\vec{r}_e + m_p\vec{r}_p}{m_e + m_p}, \qquad \vec{r} = \vec{r}_e - \vec{r}_p. \tag{6.116}$$

We can verify that the Laplacians ∇_e^2 and ∇_p^2 are related to

$$\nabla_R^2 = \frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}, \qquad \nabla_r^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{6.117}$$

as follows:

$$\frac{1}{m_e}\nabla_e^2 + \frac{1}{m_p}\nabla_p^2 = \frac{1}{M}\nabla_R^2 + \frac{1}{\mu}\nabla_r^2, \tag{6.118}$$

where

$$M = m_e + m_p, \qquad \mu = \frac{m_e m_p}{m_e + m_p} \tag{6.119}$$

are the total and reduced masses, respectively The time-independent Schrödinger equation (6.115) then becomes

$$\left[-\frac{\hbar^2}{2M}\nabla_R^2 - \frac{\hbar^2}{2\mu}\nabla_r^2 + V(r)\right]\Psi_E(\vec{R},\,\vec{r}) = E\,\Psi_E(\vec{R},\vec{r}), \tag{6.120}$$

where $\Psi_E(\vec{R},\,\vec{r}) = \chi(\vec{r}_e,\,\vec{r}_p)$. Let us now solve this equation by the separation of variables; that is, we look for solutions of the form

$$\Psi_E(\vec{R},\,\vec{r}) = \Phi(\vec{R})\psi(\vec{r}), \tag{6.121}$$

where $\Phi(\vec{R})$ and $\psi(\vec{r})$ are the wave functions of the CM and of the relative motions, respectively. Substituting this wave function into (6.120) and dividing by $\Phi(\vec{R})\psi(\vec{r})$, we obtain

$$\left[-\frac{\hbar^2}{2M}\frac{1}{\Phi(\vec{R})}\nabla_R^2\Phi(\vec{R})\right] + \left[-\frac{\hbar^2}{2\mu}\frac{1}{\psi(\vec{r})}\nabla_r^2\psi(\vec{r}) + V(r)\right] = E. \tag{6.122}$$

The first bracket depends only on $\vec{R}$ whereas the second bracket depends only on $\vec{r}$. Since $\vec{R}$ and $\vec{r}$ are independent vectors, the two expressions of the left hand side of (6.122) must be separately constant. Thus, we can reduce (6.122) to the following two separate equations:

$$-\frac{\hbar^2}{2M}\nabla_R^2\Phi(\vec{R}) = E_R\Phi(\vec{R}), \tag{6.123}$$

$$-\frac{\hbar^2}{2\mu}\nabla_r^2\psi(\vec{r}) + V(r)\psi(\vec{r}) = E_r\psi(\vec{r}), \tag{6.124}$$

with the condition

$$E_R + E_r = E. \tag{6.125}$$

We have thus reduced the Schrödinger equation (6.120), which involves two variables $\vec{R}$ and $\vec{r}$, into two separate equations (6.123) and (6.124) each involving a single variable. Note that equation (6.123) shows that the center of mass moves like a free particle of mass M. The solution to this kind of equation was examined earlier in this chapter; it has the form

$$\Phi(\vec{R}) = (2\pi)^{-3/2}e^{i\vec{k}\cdot\vec{R}}, \tag{6.126}$$

where $\vec{k}$ is the wave vector associated with the center of mass. The constant $E_R = \hbar^2k^2/(2M)$ gives the kinetic energy of the center of mass in the lab system (the total mass M is located at the origin of the center of mass coordinate system).

The second equation (6.124) represents the Schrödinger equation of a fictitious particle of mass μ moving in the central potential $-e^2/r$.

We should note that the total wave function $\Psi_E(\vec{R},\vec{r}) = \Phi(\vec{R})\psi(\vec{r})$ is seldom used. When the hydrogen problem is mentioned, this implicitly refers to $\psi(\vec{r})$ and E_r. That is, the hydrogen wave function and energy are taken to be given by $\psi(\vec{r})$ and E_r, not by Ψ_E and E.

6.3.5.2 Solution of the Radial Equation for the Hydrogen Atom

The Schrödinger equation (6.124) for the relative motion has the form of an equation for a central potential. The wave function $\psi(\vec{r})$ that is a solution to this equation is a product of an angular part and a radial part. The angular part is given by the spherical harmonic $Y_{lm}(\theta,\varphi)$. The radial part $R(r)$ can be obtained by solving the following radial equation:

$$-\frac{\hbar^2}{2\mu}\frac{d^2U(r)}{dr^2}+\left[\frac{l(l+1)\hbar^2}{2\mu r^2}-\frac{e^2}{r}\right]U(r)=EU(r), \tag{6.127}$$

where $U(r)=rR(r)$. To solve this radial equation, we are going to consider first its asymptotic solutions and then attempt a power series solution.

(a) Asymptotic behavior of the radial wave function
For very small values of r, (6.127) reduces to

$$-\frac{d^2U(r)}{dr^2}+\frac{l(l+1)}{r^2}U(r)=0, \tag{6.128}$$

whose solutions are of the form

$$U(r)=Ar^{l+1}+Br^{-l}, \tag{6.129}$$

where A and B are constants. Since $U(r)$ vanishes at $r=0$, the second term r^{-l}, which diverges at $r=0$, must be discarded. Thus, for small r, the solution is

$$U(r)\sim r^{l+1}. \tag{6.130}$$

Now, in the limit of very large values of r, we can approximate (6.127) by

$$\frac{d^2U(r)}{dr^2}+\frac{2\mu E}{\hbar^2}U(r)=0. \tag{6.131}$$

Note that, for bound state solutions, which correspond to the states where the electron and the proton are bound together, the energy E must be negative. Hence the solutions to this equation are of the form $U(r)\sim e^{\pm\lambda r}$ where $\lambda=\sqrt{2\mu(-E)}/\hbar$. Only the minus sign solution is physically acceptable, since $e^{\lambda r}$ diverges for large values of r. So, for large values of r, $U(r)$ behaves like

$$U(r)\longrightarrow e^{-\lambda r}. \tag{6.132}$$

The solutions to (6.127) can be obtained by combining (6.130) and (6.132):

$$U(r)=r^{l+1}f(r)e^{-\lambda r}, \tag{6.133}$$

where $f(r)$ is an r-dependent function. Substituting (6.133) into (6.127) we end up with a differential equation that determines the form of $f(r)$:

$$\frac{d^2f}{dr^2}+2\left(\frac{l+1}{r}-\lambda\right)\frac{df}{dr}+2\left[\frac{-\lambda(l+1)+\mu e^2/\hbar^2}{r}\right]f(r)=0. \tag{6.134}$$

(b) Power series solutions for the radial equation

As in the case of the three-dimensional harmonic oscillator, let us try a power series solution for (6.134):

$$f(r) = \sum_{k=0}^{\infty} b_k r^k, \tag{6.135}$$

which, when inserted into (6.134), yields

$$\sum_{k=0}^{\infty} \left\{ k(k+2l+1) b_k r^{k-2} + 2\left[-\lambda(k+l+1) + \frac{\mu e^2}{\hbar^2} \right] b_k r^{k-1} \right\} = 0. \tag{6.136}$$

This equation leads to the following recurrence relation (by changing k to $k-1$ in the last term):

$$k(k+2l+1) b_k = 2\left[\lambda(k+l) - \frac{\mu e^2}{\hbar^2} \right] b_{k-1}. \tag{6.137}$$

In the limit of large values of k, the ratio of successive coefficients,

$$\frac{b_k}{b_{k-1}} = \frac{2\left[\lambda(k+l) - \mu e^2/\hbar^2\right]}{k(k+2l+1)}, \tag{6.138}$$

is of the order of

$$\frac{b_k}{b_{k-1}} \longrightarrow \frac{2\lambda}{k}. \tag{6.139}$$

This is the behavior of an exponential series, since the ratio of successive coefficients of the relation $e^{2x} = \sum_{k=0}^{\infty} (2x)^k/k!$ is given by

$$\frac{2^k}{k!} \frac{(k-1)!}{2^{k-1}} = \frac{2}{k}. \tag{6.140}$$

That is, the asymptotic behavior of (6.135) is

$$f(r) = \sum_{k=0}^{\infty} b_k r^k \longrightarrow e^{2\lambda r}; \tag{6.141}$$

hence the radial solution (6.133) becomes

$$U(r) = r^{l+1} e^{2\lambda r} e^{-\lambda r} = r^{l+1} e^{\lambda r}. \tag{6.142}$$

But this contradicts (6.133): for large values of r, the asymptotic behavior of the physically acceptable radial function (6.133) is given by $e^{-\lambda r}$ while that of (6.142) by $e^{\lambda r}$; the form (6.142) is thus physically unacceptable.

(c) Energy quantization

To obtain physically acceptable solutions, the series (6.135) *must terminate at a certain power N*; hence the function $f(r)$ becomes a *polynomial* of order N:

$$f(r) = \sum_{k=0}^{N} b_k r^k. \tag{6.143}$$

This requires that all coefficients b_{N+1}, b_{N+2}, b_{N+3}, ... have to vanish. When $b_{N+1} = 0$ the recurrence formula (6.137) yields

$$\lambda(N + l + 1) - \frac{\mu e^2}{\hbar^2} = 0. \tag{6.144}$$

Since $\lambda = \sqrt{-2\mu E/\hbar^2}$ and using the notation

$$n = N + l + 1, \tag{6.145}$$

where n is known as the *principal* quantum number and N as the *radial* quantum number, we can infer the energy

$$E_n = -\frac{\mu e^4}{2\hbar^2}\frac{1}{n^2}, \tag{6.146}$$

which in turn can be written as

$$\boxed{E_n = -\frac{\mu e^4}{2\hbar^2}\frac{1}{n^2} = -\frac{e^2}{2a_0}\frac{1}{n^2},} \tag{6.147}$$

because (from Bohr theory of the hydrogen atom) the Bohr radius is given by $a_0 = \hbar^2/(\mu e^2)$ and hence $\mu/\hbar^2 = 1/(e^2 a_0)$. Note that we can write λ in terms of a_0 as follows:

$$\lambda = \sqrt{-2\frac{\mu}{\hbar^2}E_n} = \sqrt{2\frac{1}{e^2 a_0}\frac{e^2}{2a_0 n^2}} = \frac{1}{na_0}. \tag{6.148}$$

Since $N = 0, 1, 2, 3, \ldots$, the allowed values of n are nonzero integers, $n = l + 1, l + 2, l + 3, \ldots$. For a given value of n, the orbital quantum number l can have values only between 0 and $n - 1$ (i.e., $l = 0, 1, 2, \ldots, n - 1$).

Remarks

- Note that (6.147) is similar to the energy expression obtained from the Bohr quantization condition, discussed in Chapter 1. It can be rewritten in terms of the Rydberg constant $\mathcal{R} = m_e e^4/(2\hbar^2)$ as follows:

$$E_n = -\frac{m_p}{m_p + m_e}\frac{\mathcal{R}}{n^2}, \tag{6.149}$$

  where $\mathcal{R} = 13.6$ eV. Since the ratio m_e/m_p is very small ($m_e/m_p \ll 1$), we can approximate this expression by

$$E_n = -\left(1 + \frac{m_e}{m_p}\right)^{-1}\frac{\mathcal{R}}{n^2} \simeq -\left(1 - \frac{m_e}{m_p}\right)\frac{\mathcal{R}}{n^2}. \tag{6.150}$$

  So, if we consider the proton to be infinitely more massive than the electron, we recover the energy expression as derived by Bohr: $E_n = -\mathcal{R}/n^2$.

- **Energy of hydrogen-like atoms:** How does one obtain the energy of an atom or ion with a nuclear charge Ze but which has only one electron[4]? Since the Coulomb potential

[4]For instance, $Z = 1$ refers to H, $Z = 2$ to He^+, $Z = 3$ to Li^{2+}, $Z = 4$ to Be^{3+}, $Z = 5$ to B^{4+}, $Z = 6$ to C^{5+}, and so on.

felt by the single electron due to the charge Ze is given by $V(r) = -Ze^2/r$, the energy of the electron can be inferred from (6.147) by simply replacing e^2 with Ze^2:

$$E_n = -\frac{m_e(Ze^2)^2}{2\hbar^2}\frac{1}{n^2} = -\frac{Z^2E_0}{n^2}, \tag{6.151}$$

where $E_0 = e^2/(2a_0) = 13.6\,\text{eV}$; in deriving this relation, we have assumed that the mass of the nucleus is infinitely large compared to the electronic mass.

(d) Radial wave functions of the hydrogen atom

The radial wave function $R_{nl}(r)$ can be obtained by inserting (6.143) into (6.133),

$$R_{nl}(r) = \frac{1}{r}U_{nl}(r) = A_{nl}r^l e^{-\lambda r}\sum_{k=0}^{N} b_k r^k = A_{nl}r^l e^{-r/na_0}\sum_{k=0}^{N} b_k r^k, \tag{6.152}$$

since, as shown in (6.148), $\lambda = 1/(na_0)$; A_{nl} is a normalization constant.

How does one determine the expression of $R_{nl}(r)$? This issue reduces to obtaining the form of the polynomial $r^l\sum_{k=0}^{N} b_k r^k$ and the normalization constant A_{nl}. For this, we are going to explore two methods: the first approach follows a straightforward calculation and the second makes use of special functions.

(i) First approach: straightforward calculation of $R_{nl}(r)$

This approach consists of a straightforward construction of $R_{nl}(r)$; we are going to show how to construct only the first few expressions. For instance, if $n = 1$ and $l = 0$ then $N = 0$. Since $N = n - l - 1$ and $\lambda = 1/(na_0)$ we can write (6.152) as

$$R_{10}(r) = A_{10}e^{-r/a_0}\sum_{k=0}^{0} b_k r^k = A_{10}b_0 e^{-r/a_0}, \tag{6.153}$$

where $A_{10}b_0$ can be obtained from the normalization of $R_{10}(r)$: using $\int_0^\infty x^n e^{-ax}\,dx = n!/a^{n+1}$, we have

$$1 = \int_0^\infty r^2 |R_{10}(r)|^2\,dr = A_{10}^2 b_0^2 \int_0^\infty r^2 e^{-2r/a_0} dr = A_{10}^2 b_0^2 \frac{a_0^3}{4}; \tag{6.154}$$

hence $A_{10} = 1$ and $b_0 = 2\,(a_0)^{-3/2}$. Thus, $R_{10}(r)$ is given by

$$R_{10}(r) = 2\,(a_0)^{-3/2}\,e^{-r/a_0}. \tag{6.155}$$

Next, let us find $R_{20}(r)$. Since $n = 2$, $l = 0$ we have $N = 2 - 0 - 1 = 1$ and

$$R_{20}(r) = A_{20}e^{-r/2a_0}\sum_{k=0}^{1} b_k r^k = A_{20}(b_0 + b_1 r)e^{-r/2a_0}. \tag{6.156}$$

From (6.138) we can express b_1 in terms of b_0 as

$$b_1 = \frac{2\lambda(k+l) - 2/a_0}{k(k+2l+1)}b_0 = -\frac{1}{2a_0}b_0 = -\frac{1}{a_0\sqrt{a_0^3}}, \tag{6.157}$$

because $\lambda = 1/(2a_0)$, $k = 1$, and $l = 0$. So, substituting (6.157) into (6.156) and normalizing, we get $A_{20} = 1/(2\sqrt{2})$; hence

$$R_{20}(r) = \frac{1}{\sqrt{2a_0^3}}\left(1 - \frac{r}{2a_0}\right)e^{-r/2a_0}. \tag{6.158}$$

Continuing in this way, we can obtain the expression of any radial wave function $R_{nl}(r)$; note that, knowing $b_0 = 2\,(a_0)^{-3/2}$, we can use the recursion relation (6.138) to obtain all other coefficients $b_2, b_3, \ldots$.

(ii) Second approach: determination of $R_{nl}(r)$ by means of special functions

The polynomial $r^l \sum_{k=0}^{N} b_k r^k$ present in (6.152) is a polynomial of degree $N + l$ or $n - 1$ since $n = N + l + 1$. This polynomial, which is denoted by $L_k^N(r)$, is known as the *associated Laguerre polynomial*; it is a solution to the Schrödinger equation (6.134). The solutions to differential equations of the form (6.134) were studied by Laguerre long before the birth of quantum mechanics. The associated Laguerre polynomial is defined, in terms of the *Laguerre polynomials* of order k, $L_k(r)$, by

$$L_k^N(r) = \frac{d^N}{dr^N}L_k(r), \tag{6.159}$$

where

$$L_k(r) = e^r \frac{d^k}{dr^k}(r^k e^{-r}). \tag{6.160}$$

Table 6.5: First few Laguerre polynomials and associated Laguerre polynomials.

Laguerre polynomials $L_k(r)$	Associated Laguerre polynomials $L_k^N(r)$
$L_0 = 1$	
$L_1 = 1 - r$	$L_1^1 = -1$
$L_2 = 2 - 4r + r^2$	$L_2^1 = -4 + 2r, \quad L_2^2 = 2$
$L_3 = 6 - 18r + 9r^2 - r^3$	$L_3^1 = -18 + 18r - 3r^2,\ L_3^2 = 18 - 6r,\ L_3^3 = -6$
$L_4 = 24 - 96r + 72r^2 - 16r^3 + r^4$	$L_4^1 = -96 + 144r - 48r^2 + 4r^3$
	$L_4^2 = 144 - 96r + 12r^2,\ L_4^3 = 24r - 96,\ L_4^4 = 24$
$L_5 = 120 - 600r + 600r^2 - 200r^3$	$L_5^1 = -600 + 1200r - 600r^2 + 100r^3 - 5r^4$
$+ 25r^4 - r^5$	$L_5^2 = 1200 - 1200r + 300r^2 - 20r^3$
	$L_5^3 = -1200 + 600r - 60r^2,\ L_5^4 = 600 - 120r$
	$L_5^5 = -120$

The first few Laguerre polynomials are listed in Table 6.5.

We can verify that $L_k(r)$ and $L_k^N(r)$ satisfy the following differential equations:

$$r\frac{d^2L_k(r)}{dr^2} + (1-r)\frac{dL_k(r)}{dr} + kL_k(r) = 0, \tag{6.161}$$

$$r\frac{d^2L_k^N(r)}{dr^2} + (N+1-r)\frac{dL_k^N(r)}{dr} + (k-N)L_k^N(r) = 0. \tag{6.162}$$

This last equation is identical to the hydrogen atom radial equation (6.134). The proof goes as follows. Using a change of variable

$$\rho = 2\lambda r = 2\frac{\sqrt{-2\mu E}}{\hbar}r, \tag{6.163}$$

along with the fact that $a_0 = \hbar^2/(\mu e^2)$ (Bohr radius), we can show that (6.134) reduces to

$$\rho\frac{d^2g(\rho)}{d\rho^2} + [(2l+1)+1-\rho]\frac{dg(\rho)}{d\rho} + [(n+l)-(2l+1)]\,g(\rho) = 0, \tag{6.164}$$

where $f(r) = g(\rho)$. In deriving (6.164), we have used the fact that $1/\lambda a_0 = n$ (see (6.148)). Note that equations (6.162) and (6.164) are identical; the solutions to (6.134) are thus given by the associated Laguerre polynomials $L_{n+l}^{2l+1}(2\lambda r)$.

The radial wave function of the hydrogen atom is then given by

$$R_{nl}(r) = N_{nl}\left(\frac{2r}{na_0}\right)^l e^{-r/na_0}L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right), \tag{6.165}$$

where N_{nl} is a constant obtained by normalizing the radial function $R_{nl}(r)$:

$$\int_0^\infty r^2R_{nl}^2(r)\,dr = 1. \tag{6.166}$$

Table 6.6: The first few radial wave functions $R_{nl}(r)$ of the hydrogen atom.

$R_{10}(r) = 2a_0^{-3/2}e^{-r/a_0}$	$R_{21}(r) = \frac{1}{\sqrt{6a_0^3}}\frac{r}{2a_0}e^{-r/2a_0}$
$R_{20}(r) = \frac{1}{\sqrt{2a_0^3}}\left(1-\frac{r}{2a_0}\right)e^{-r/2a_0}$	$R_{31}(r) = \frac{8}{9\sqrt{6a_0^3}}\left(1-\frac{r}{6a_0}\right)\left(\frac{r}{3a_0}\right)e^{-r/3a_0}$
$R_{30}(r) = \frac{2}{\sqrt{27a_0^3}}\left(1-\frac{2r}{3a_0}+\frac{2r^2}{27a_0^2}\right)e^{-r/3a_0}$	$R_{32}(r) = \frac{4}{9\sqrt{30a_0^3}}\left(\frac{r}{3a_0}\right)^2 e^{-r/3a_0}$
$R_{40}(r) = \frac{1}{\sqrt{16a_0^3}}\left(1-\frac{3r}{4a_0}+\frac{r^2}{8a_0^2}-\frac{r^3}{192a_0^3}\right)e^{-r/4a_0}$	$R_{41}(r) = \sqrt{\frac{5}{3a_0^3}}\left(\frac{r}{16a_0}\right)\left(1-\frac{r}{4a_0}+\frac{r^2}{80a_0^2}\right)e^{-r/4a_0}$

Using the normalization condition of the associated Laguerre functions

$$\int_0^\infty e^{-\rho}\rho^{2l}\left[L_{n+l}^{2l+1}(\rho)\right]^2\rho^2 d\rho = \frac{2n\left[(n+l)!\right]^3}{(n-l-1)!}, \tag{6.167}$$

where $\rho = 2\lambda r = 2r/(na_0)$, we can show that N_{nl} is given by

$$N_{nl} = -\left(\frac{2}{na_0}\right)^{3/2}\sqrt{\frac{(n-l-1)!}{2n[(n+l)!]^3}}. \tag{6.168}$$

The wave functions of the hydrogen atom are given by

$$\boxed{\psi_{nlm}(r,\theta,\varphi) = R_{nl}(r)Y_{lm}(\theta,\varphi),} \tag{6.169}$$

where the radial functions $R_{nl}(r)$ are

$$\boxed{R_{nl}(r) = -\left(\frac{2}{na_0}\right)^{3/2}\sqrt{\frac{(n-l-1)!}{2n[(n+l)!]^3}}\left(\frac{2r}{na_0}\right)^l e^{-r/na_0}L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right).} \tag{6.170}$$

The first few radial wave functions are listed in Table 6.6; as shown in (6.155) and (6.158), they are identical with those obtained from a straightforward construction of $R_{nl}(r)$. The shapes of some of these radial functions are plotted in Figure 6.3.

(e) Properties of the radial wave functions of hydrogen
The radial wave functions of the hydrogen atom behave as follows (see Figure 6.3):

- They behave like r^l for small r.
- They decrease exponentially at large r, since L_{n+l}^{2l+1} is dominated by the highest power, r^{n-l-1}.
- Each function $R_{nl}(r)$ has $n-l-1$ radial nodes, since $L_{n+l}^{2l+1}(\rho)$ is a polynomial of degree $n-l-1$.

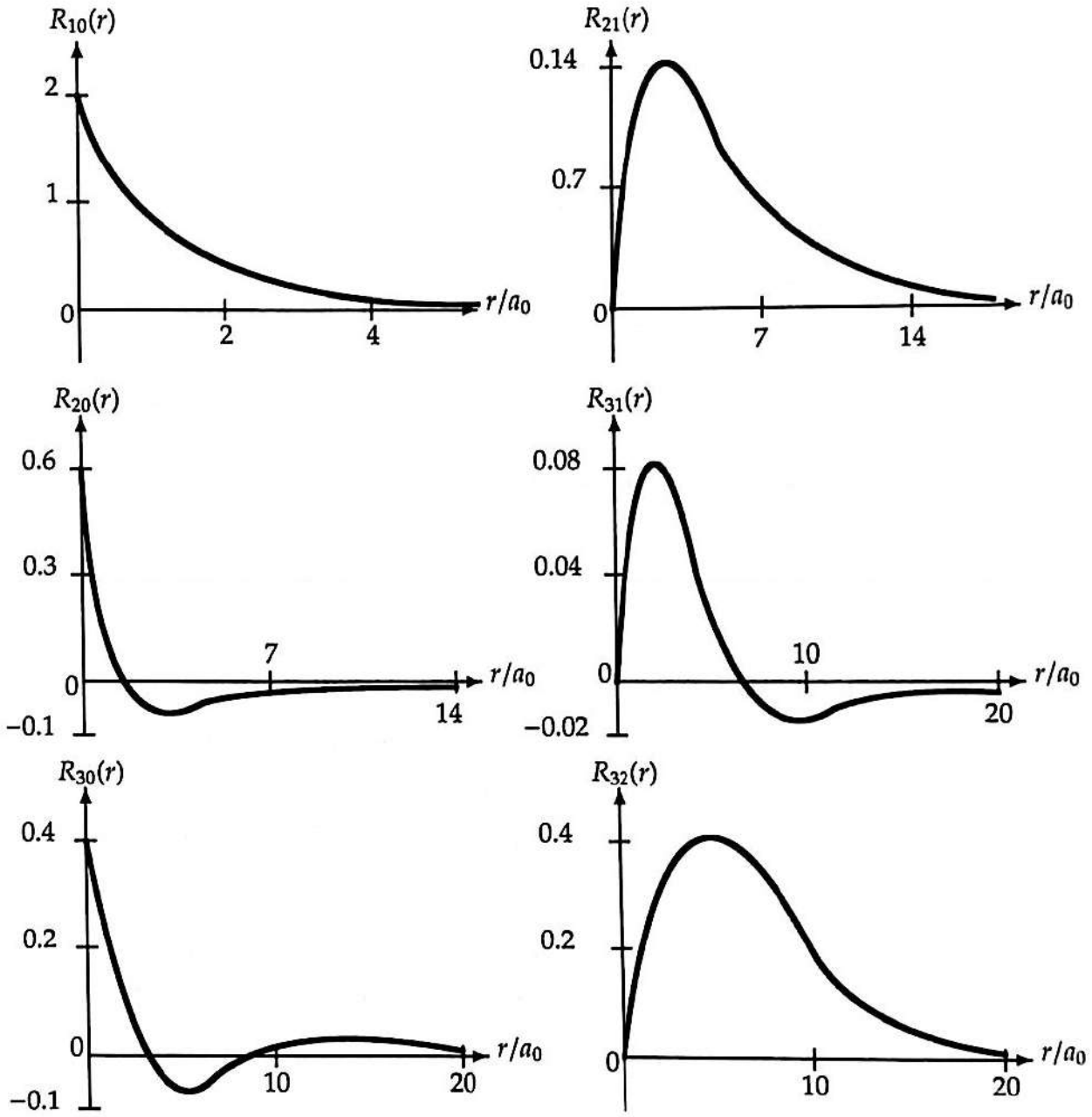


Figure 6.3: The first few radial wave functions $R_{nl}(r)$ for hydrogen; the radial length is in units of the Bohr radius $a_0 = \hbar^2/(\mu e^2)$. Notice that $R_{nl}(r)$ has $(n - l - 1)$ nodes.

Table 6.7: Hydrogen energy levels and their degeneracies when the electron's spin is ignored.

n	l	Orbitals	m	g_n	E_n
1	0	s	0	1	$-e^2/(2a_0)$
2	0	s	0	4	$-e^2/(8a_0)$
	1	p	$-1,0,1$		
3	0	s	0	9	$-e^2/(18a_0)$
	1	p	$-1,0,1$		
	2	d	$-2,-1,0,1,2$		
4	0	s	0	16	$-e^2/(32a_0)$
	1	p	$-1,0,1$		
	2	d	$-2,-1,0,1,2$		
	3	f	$-3,-2,-1,0,1,2,3$		
5	0	s	0	25	$-e^2/(50a_0)$
	1	p	$-1,0,1$		
	2	d	$-2,-1,0,1,2$		
	3	f	$-3,-2,-1,0,1,2,3$		
	4	g	$-4,-3,-2,-1,0,1,2,3,4$		

6.3.5.3 Degeneracy of the Bound States of Hydrogen

Besides being independent of m, which is a property of central potentials (see (6.55)), the energy levels (6.147) are also independent of l. This *additional degeneracy* in l is not a property of central potentials, but a *particular feature of the Coulomb potential.* In the case of central potentials, the energy E usually depends on two quantum numbers: one radial, n, and the other orbital, l, giving E_{nl}.

The total quantum number n takes only nonzero values $1, 2, 3, \ldots$. As displayed in Table 6.7, for a given n, the quantum l number may vary from 0 to $n-1$; and for each l, m can take $(2l+1)$ values: $m = -l, -l+1, \ldots, l-1, l$. The degeneracy of the state n, which is specified by the total number of different states associated with n, is then given by (see Example 6.3 on page 399)

$$\boxed{g_n = \sum_{l=0}^{n-1}(2l+1) = n^2.} \tag{6.171}$$

Remarks

1. The state of every hydrogenic electron is specified by three quantum numbers (n, l, m), called the single-particle state or *orbital*, $|nlm\rangle$. According to the spectroscopic notation, the states corresponding to the respective numerical values $l = 0, 1, 2, 3, 4, 5, \ldots$ are called the s, p, d, f, g, h, ... states; the letters s, p, d, f refer to sharp, principal, diffuse, and fundamental labels, respectively (as for the letters g, h, ..., they have yet to be assigned

labels, the reader is free to guess how to refer to them!). Hence, as shown in Table 6.7, for a given n, an s-state has 1 orbital $|n00\rangle$, a p-state has 3 orbitals $|n1m\rangle$ corresponding to $m = -1, 0, 1$, a d-state has 5 orbitals $|n2m\rangle$ corresponding to $m = -2, -1, 0, 1, 2$, and so on.

2. If we take into account the spin of the electron, the state of every electron will be specified by four quantum numbers (n, l, m_l, m_s), where $m_s = \pm\frac{1}{2}$ is the z-component of the spin of the electron. Hence the complete wave function of the hydrogen atom must be equal to the product of a space part or orbital $\psi_{nlm_l}(r, \theta, \varphi) = R_{nl}(r)Y_{lm_l}(\theta, \varphi)$, and a spin part $\left|\frac{1}{2}, m_s\right\rangle$:

$$\Psi_{nlm_lm_s}(\vec{r}) = \psi_{nlm_l}(r, \theta, \varphi)\left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle = R_{nl}(r)Y_{lm_l}(\theta, \varphi)\left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle. \tag{6.172}$$

Using the spinors from Chapter 5 we can write the spin-up wave function as

$$\Psi_{nlm_l\,\frac{1}{2}}(\vec{r}) = \psi_{nlm_l}(r, \theta, \varphi)\left|\frac{1}{2}, \frac{1}{2}\right\rangle = \psi_{nlm_l}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \psi_{nlm_l} \\ 0 \end{pmatrix}, \tag{6.173}$$

and the spin-down wave function as

$$\Psi_{nlm_l\,-\frac{1}{2}}(\vec{r}) = \psi_{nlm_l}(r, \theta, \varphi)\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \psi_{nlm_l}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \psi_{nlm_l} \end{pmatrix}. \tag{6.174}$$

For instance, the spin-up and spin-down *ground state* wave functions of hydrogen are given by

$$\Psi_{100\,\frac{1}{2}}(\vec{r}) = \begin{pmatrix} \psi_{100} \\ 0 \end{pmatrix} = \begin{pmatrix} (1/\sqrt{\pi})a_0^{-3/2}e^{-r/a_0} \\ 0 \end{pmatrix}, \tag{6.175}$$

$$\Psi_{100\,-\frac{1}{2}}(\vec{r}) = \begin{pmatrix} 0 \\ \psi_{100} \end{pmatrix} = \begin{pmatrix} 0 \\ (1/\sqrt{\pi})a_0^{-3/2}e^{-r/a_0} \end{pmatrix}. \tag{6.176}$$

3. When spin is included, the degeneracy of the hydrogen's energy levels is given by

$$2\sum_{l=0}^{n-1}(2l + 1) = 2n^2, \tag{6.177}$$

since, in addition to the degeneracy (6.171), each level is doubly degenerate with respect to the spin degree of freedom. For instance, the ground state of hydrogen is doubly degenerate since $\Psi_{100\,\frac{1}{2}}(\vec{r})$ and $\Psi_{100\,-\frac{1}{2}}(\vec{r})$ correspond to the same energy $-13.6\,\text{eV}$. Similarly, the first excited state is eightfold degenerate ($2(2)^2 = 8$) because the eight states $\Psi_{200\,\pm\frac{1}{2}}(\vec{r})$, $\Psi_{211\,\pm\frac{1}{2}}(\vec{r})$, $\Psi_{210\,\pm\frac{1}{2}}(\vec{r})$, and $\Psi_{21-1\,\pm\frac{1}{2}}(\vec{r})$ correspond to the same energy $-13.6/4\,\text{eV} = -3.4\,\text{eV}$.

6.3.5.4 Probabilities and Averages

When a hydrogen atom is in the stationary state $\psi_{nlm}(r, \theta, \varphi)$, the quantity $|\psi_{nlm}(r, \theta, \varphi)|^2\, d^3r$ represents the probability of finding the electron in the volume element d^3r, where

$d^3r = r^2 \sin\theta\, dr\, d\theta\, d\varphi$. The probability of finding the electron in a spherical shell located between r and $r + dr$ (i.e., a shell of thickness dr) is given by

$$\begin{aligned} P_{nl}(r)\,dr &= \left(\int_0^{\pi} \sin\theta\, d\theta \int_0^{2\pi} d\varphi\, |\psi_{nlm}(r,\theta,\varphi)|^2\right) r^2 dr \\ &= |R_{nl}(r)|^2 r^2 dr \int_0^{\pi} \sin\theta\, d\theta \int_0^{2\pi} Y^*_{lm}(\theta,\varphi) Y_{lm}(\theta,\varphi)\, d\varphi \\ &= |R_{nl}(r)|^2 r^2 dr. \end{aligned} \tag{6.178}$$

If we integrate this quantity between $r = 0$ and $r = a$, we obtain the probability of finding the electron in a sphere of radius a centered about the origin. Hence integrating between $r = 0$ and $r = \infty$, we would obtain 1, which is the probability of finding the electron somewhere in space.

Let us now specify the average values of the various powers of r. Since $\psi_{nlm}(r,\theta,\varphi) = R_{nl}(r)Y_{lm}(\theta,\varphi)$, we can see that the average of r^k is independent of the azimuthal quantum number m:

$$\begin{aligned} \langle nlm|r^k|nlm\rangle &= \int r^k |\psi_{nlm}(r,\theta,\varphi)|^2 r^2 \sin\theta\, dr\, d\theta\, d\varphi \\ &= \int_0^{\infty} r^{k+2} |R_{nl}(r)|^2 dr \int_0^{\pi} \sin\theta\, d\theta \int_0^{2\pi} Y^*_{lm}(\theta,\varphi) Y_{lm}(\theta,\varphi)\, d\varphi \\ &= \int_0^{\infty} r^{k+2} |R_{nl}(r)|^2 dr \\ &= \langle nl \mid r^k | nl\rangle. \end{aligned} \tag{6.179}$$

Using the properties of Laguerre polynomials, we can show that (Problem 6.2, page 412)

$$\langle nl|\, r\, |nl\rangle = \frac{1}{2}\left[3n^2 - l(l+1)\right] a_0, \tag{6.180}$$

$$\langle nl|\, r^2 \mid nl\rangle = \frac{1}{2} n^2 \left[5n^2 + 1 - 3l\,(l+1)\right] a_0^2, \tag{6.181}$$

$$\langle nl|r^{-1}|nl\rangle = \frac{1}{n^2\, a_0}, \tag{6.182}$$

$$\langle nl|r^{-2}|nl\rangle = \frac{2}{n^3(2l\ +1)a_0^2}, \tag{6.183}$$

where a_0 is the Bohr radius, $a_0 = \hbar^2/(\mu e^2)$. The averages (6.180) to (6.183) can be easily derived from Kramers' recursion relation (Problem 6.3, page 413):

$$\frac{k+1}{n^2}\langle nl|r^k|nl\rangle - (2k+1)a_0\langle nl|r^{k-1}|nl\rangle + \frac{ka_0^2}{4}\left[(2l+1)^2 - k^2\right]\langle nl|r^{k-2}|nl\rangle = 0. \tag{6.184}$$

Equations (6.180) and (6.182) reveal that $1/\langle r\rangle$ and $\langle 1/r\rangle$ are not equal, but are of the same order of magnitude:

$$\langle r\rangle \sim n^2 a_0. \tag{6.185}$$

This relation is in agreement with the expression obtained from the Bohr theory of hydrogen: the quantized radii of circular orbits for the hydrogen atom are given by $r_n = n^2 a_0$. We will show in Problem 6.6 page 417 that the Bohr radii for circular orbits give the locations where the probability density of finding the electron reaches its maximum.

Next, using the expression (6.182) for $\langle r^{-1} \rangle$, we can obtain the average value of the Coulomb potential

$$\langle V(r) \rangle = -e^2 \left\langle \frac{1}{r} \right\rangle = -\frac{e^2}{a_0} \frac{1}{n^2}, \tag{6.186}$$

which, as specified by (6.147), is equal to twice the total energy:

$$E_n = \frac{1}{2} \langle V(r) \rangle = -\frac{e^2}{2a_0} \frac{1}{n^2}. \tag{6.187}$$

This is known as the Virial theorem, which states that if $V(\alpha r) = \alpha^n V(r)$, the average expressions of the kinetic and potential energies are related by

$$\langle T \rangle = \frac{n}{2} \langle V(r) \rangle. \tag{6.188}$$

For instance, in the case of a Coulomb potential $V(\alpha r) = \alpha^{-1} V(r)$, we have $\langle T \rangle = -\frac{1}{2} \langle V \rangle$; hence $E = -\frac{1}{2} \langle V \rangle + \langle V \rangle = \frac{1}{2} \langle V \rangle$.

Example 6.3 (Degeneracy relation for the hydrogen atom)
Prove the degeneracy relation (6.171) for the hydrogen atom.

Solution
The energy $E_n = -e^2/(2a_0 n^2)$ of the hydrogen atom (6.147) does not depend on the orbital quantum number l or on the azimuthal number m; it depends only on the principal quantum number n. For a given n, the orbital number l can take $n-1$ values: $l = 0, 1, 2, 3, \ldots, n-1$; while for each l, the azimuthal number m takes $2l+1$ values: $m = -l, -l+1, \ldots, l-1, l$. Thus, for each n, there exist g_n different wave functions $\psi_{nlm}(\vec{r})$, which correspond to the same energy E_n, with

$$g_n = \sum_{l=0}^{n-1} (2l+1) = 2 \sum_{l=0}^{n-1} l + \sum_{l=0}^{n-1} 1 = n(n-1) + n = n^2. \tag{6.189}$$

Another way of finding this result consists of writing $\sum_{l=0}^{n-1} (2l+1)$ in the following two equivalent forms:

$$g_n = 1 + 3 + 5 + 7 + \cdots + (2n-7) + (2n-5) + (2n-3) + (2n-1), \tag{6.190}$$

$$g_n = (2n-1) + (2n-3) + (2n-5) + (2n-7) + \cdots + 7 + 5 + 3 + 1, \tag{6.191}$$

and then add them, term by term:

$$2g_n = (2n) + (2n) + (2n) + (2n) + \cdots + (2n) + (2n) + (2n) + (2n). \tag{6.192}$$

Since there are n terms (because l can take n values: $l = 0, 1, 2, 3, \ldots, n-1$), we have $2g_n = n(2n)$; hence $g_n = n^2$.

6.3.6 Effect of Magnetic Fields on Central Potentials

As discussed earlier (6.55), the energy levels of a particle in a central potential do not depend on the azimuthal quantum number m. This degeneracy can be lifted if we place the particle in a uniform magnetic field $\vec{B}$ (if $\vec{B}$ is uniform, its spatial derivatives vanish). We should note that we will be using the SI units throughout this subsection.

6.3.6.1 Effect of a Magnetic Field on a Charged Particle

Consider a particle of mass[5] m_q and charge q which, besides moving in a central potential $V(r)$, is subject to a uniform magnetic field $\vec{B}$.

From the theory of classical electromagnetism, the vector potential corresponding to a uniform magnetic field may be written as $\vec{A} = \frac{1}{2}(\vec{B} \times \vec{r})$ since, using the relation $\vec{\nabla} \times (\vec{C} \times \vec{D}) = \vec{C}(\vec{\nabla} \cdot \vec{D}) - \vec{D}(\vec{\nabla} \cdot \vec{C}) + (\vec{D} \cdot \vec{\nabla})\vec{C} - (\vec{C} \cdot \vec{\nabla})\vec{D}$, we have

$$\vec{\nabla} \times \vec{A} = \frac{1}{2}\vec{\nabla} \times (\vec{B} \times \vec{r}) = \frac{1}{2}\left[\vec{B}(\vec{\nabla} \cdot \vec{r}) - (\vec{B} \cdot \vec{\nabla})\vec{r}\right] = \frac{1}{2}\left[3\vec{B} - \vec{B}\right] = \vec{B}, \tag{6.193}$$

where we have used the relations $\vec{\nabla} \cdot \vec{B} = 0$, $(\vec{r} \cdot \vec{\nabla})\vec{B} = 0$, $\vec{\nabla} \cdot \vec{r} = 3$, and $(\vec{B} \cdot \vec{\nabla})\vec{r} = \vec{B}$.

When the charge q is placed in a magnetic field $\vec{B}$, the interaction between the charge and the magnetic field can be implemented by means of the *minimal coupling* principle which consists in substituting[6]

$$\hat{\vec{p}} \longrightarrow \hat{\vec{p}} - q\vec{A}, \tag{6.194}$$

in the Hamiltonian for a free particle (see (6.124)). The Hamiltonian of the particle of charge q in the external magnetic field is thus given by[7]

$$\hat{H} = \frac{1}{2m_q}\left(\hat{\vec{p}} - q\vec{A}\right)^2 + V(r) = \hat{H}_0 - \frac{q}{2m_q}\left(\hat{\vec{p}} \cdot \vec{A} + \vec{A} \cdot \hat{\vec{p}}\right) + \frac{q^2}{2m_q}\vec{A}^2, \tag{6.195}$$

where $\hat{H}_0 = \hat{\vec{p}}^{\,2}/(2m_q) + V(r)$ is the Hamiltonian of an uncharged particle (or for a charged particle with no external magnetic field). If the *Coulomb gauge* condition $\vec{\nabla} \cdot \vec{A} = 0$ is satisfied, we can use (6.195) to write the Schrödinger equation in an electromagnetic field as follows:

$$i\hbar\frac{\partial\psi(\vec{r}, t)}{\partial t} = \left[\frac{1}{2m_q}\hat{\vec{p}}^{\,2} + \hat{V}(r) - \frac{q}{2m_q}\vec{A} \cdot \hat{\vec{p}} + \frac{q^2}{2m_q}\vec{A}^{\,2}\right]\psi(\vec{r}, t). \tag{6.196}$$

In Chapter 12 (see Eq. (12.143)), we will show that this equation comes out naturally from the non-relativistic limit of the Klein-Gordon equation.

Now, we should note that the term $\hat{\vec{p}} \cdot \vec{A}$ in (6.195) can be calculated by analogy with the commutator $[\hat{p}, \hat{F}(x)] = -i\hbar d\hat{F}(x)/dx$:

$$(\hat{\vec{p}} \cdot \vec{A}) \mid \psi\rangle = -i\hbar(\vec{\nabla} \cdot \vec{A}) \mid \psi\rangle - i\hbar\vec{A} \cdot \vec{\nabla} \mid \psi\rangle = -i\hbar(\vec{\nabla} \cdot \vec{A}) \mid \psi\rangle + \vec{A} \cdot \hat{\vec{p}} \mid \psi\rangle. \tag{6.197}$$

[5] In this section, we are going to use m_q to denote the mass of the particle of charge q so as not to create a confusion with the azimuthal quantum number m and the magnetic moment μ since they appear quite often in the equations; in future sections where we don't deal with these two quantum numbers, we will revert back to m to denote the mass.

[6] Note that Equation (6.194) is expressed in SI units. Had we used the Gaussian (or cgs) units, the minimal coupling relation (6.194) would have been written as: $\hat{\vec{p}} \longrightarrow \hat{\vec{p}} - (q/c)\vec{A}$.

[7] In Chapter 12, we provide a derivation of this Hamiltonian using the Lagrangian formalism (see Eq. (13.136)).

We see that, whenever the Coulomb gauge $\vec{\nabla}\cdot\vec{A}=0$ is satisfied, the term $\vec{A}\cdot\hat{\vec{p}}$ becomes equal to $\hat{\vec{p}}\cdot\vec{A}$:

$$\hat{\vec{p}}\cdot\vec{A}-\vec{A}\cdot\hat{\vec{p}}=-i\hbar(\vec{\nabla}\cdot\vec{A})=0 \quad\Longrightarrow\quad \vec{A}\cdot\hat{\vec{p}}=\hat{\vec{p}}\cdot\vec{A}. \tag{6.198}$$

On the other hand, since $\vec{A}=\frac{1}{2}(\vec{B}\times\vec{r})$, we have

$$\vec{A}\cdot\hat{\vec{p}}=\frac{1}{2}(\vec{B}\times\vec{r})\cdot\hat{\vec{p}}=\frac{1}{2}\vec{B}\cdot(\vec{r}\times\hat{\vec{p}})=\frac{1}{2}\vec{B}\cdot\hat{\vec{L}}, \tag{6.199}$$

where $\hat{\vec{L}}$ is the orbital angular momentum operator of the particle. Now, a combination of (6.198) and (6.199) leads to $\hat{\vec{p}}\cdot\vec{A}=\vec{A}\cdot\hat{\vec{p}}=\frac{1}{2}\vec{B}\cdot\hat{\vec{L}}$ which, when inserted in the Hamiltonian (6.195), yields

$$\hat{H}=\hat{H}_0-\frac{q}{m_q}\vec{A}\cdot\hat{\vec{p}}+\frac{q^2}{2m_q}\vec{A}^2=\hat{H}_0-\frac{q}{2m_q}\vec{B}\cdot\hat{\vec{L}}+\frac{q^2}{2m_q}\vec{A}^2=\hat{H}_0-\vec{\mu}_L\cdot\vec{B}+\frac{q^2}{2m_q}\vec{A}^2, \tag{6.200}$$

where

$$\vec{\mu}_L=\frac{q}{2m_q}\hat{\vec{L}}=\frac{\mu_B}{\hbar}\hat{\vec{L}} \tag{6.201}$$

is called the *orbital magnetic dipole moment* of the charge q and $\mu_B=q\hbar/(2m_q)$ is known as the *Bohr magneton;* for an electron, the Bohr magneton is $\mu_B=\frac{e\hbar}{2m_e}=9.274\times10^{-24}\text{J/T}$.

As mentioned in Chapter 5, $\vec{\mu}_L$ is due to the orbiting motion of the charge about the center of the potential. The term $-\vec{\mu}_L\cdot\vec{B}$ in (6.200) represents the energy resulting from the interaction between the particle's orbital magnetic dipole moment $\vec{\mu}_L=q\hat{\vec{L}}/(2m_q)$ and the magnetic field $\vec{B}$. We should note that if the charge q had an intrinsic spin $\vec{S}$, its spinning motion would give rise to a spin magnetic dipole moment $\vec{\mu}_S=q\vec{S}/(2m_q)$ which, when interacting with an external magnetic field $\vec{B}$, would in turn generate an energy term $-\vec{\mu}_S\cdot\vec{B}$ that must be added to the Hamiltonian. This issue will be discussed further in Chapter 7.

Finally, using the relation $(\vec{C}\times\vec{D})\cdot(\vec{E}\times\vec{F})=(\vec{C}\cdot\vec{E})(\vec{D}\cdot\vec{F})-(\vec{C}\cdot\vec{F})(\vec{D}\cdot\vec{E})$, and since $\vec{A}=\frac{1}{2}(\vec{B}\times\vec{r})$, we have

$$\vec{A}^2=\frac{1}{4}(\vec{B}\times\vec{r})\cdot(\vec{B}\times\vec{r})=\frac{1}{4}\left[B^2r^2-(\vec{B}\cdot\vec{r})^2\right]. \tag{6.202}$$

We can thus write (6.200) as

$$\boxed{\hat{H}=\frac{1}{2m_q}\hat{\vec{p}}^2+\hat{V}(r)-\frac{q}{2m_q}\vec{B}\cdot\hat{\vec{L}}+\frac{q^2}{8m_q}\left[B^2r^2-(\vec{B}\cdot\vec{r})^2\right].} \tag{6.203}$$

This is the Hamiltonian of a particle of mass m_q and charge q moving in a central potential $\hat{V}(r)$ under the influence of a uniform magnetic field $\vec{B}$.

6.3.6.2 The Normal Zeeman Effect ($\vec{S}=0$)

When a hydrogen atom is placed in an external uniform magnetic field, its energy levels get shifted. This energy shift is known as the *Zeeman effect*.

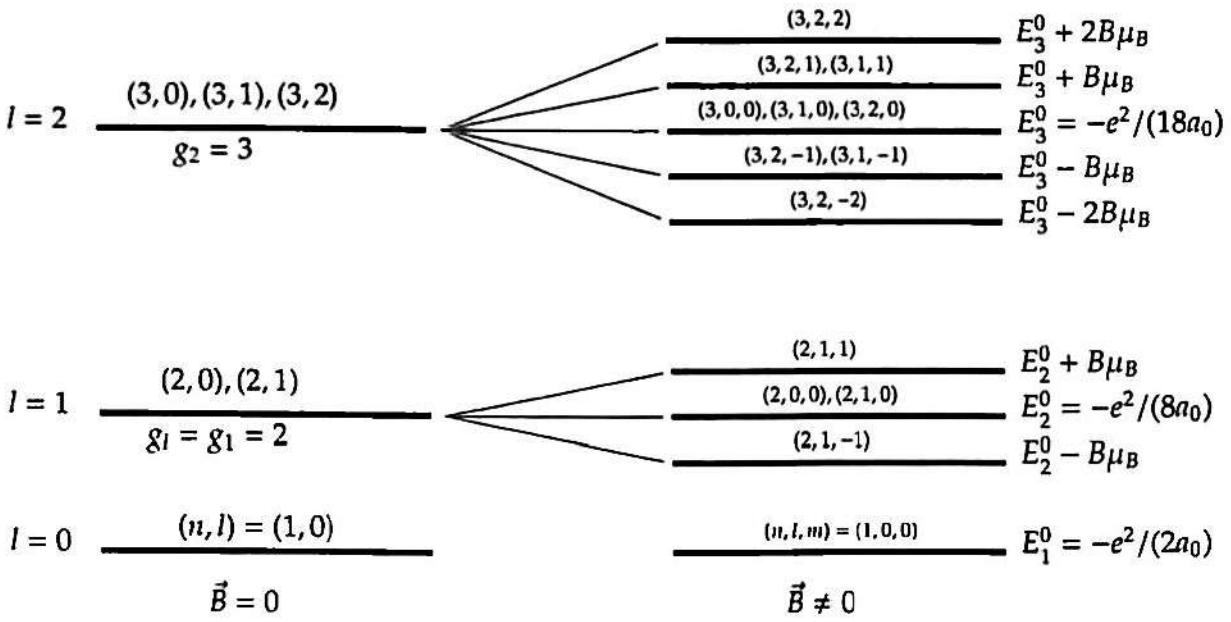


Figure 6.4: Normal Zeeman effect in hydrogen. (Left) When $\vec{B} = 0$ the energy levels are degenerate with respect to l and m. (Right) When $\vec{B} \neq 0$ the degeneracy with respect to m is removed, but the degeneracy with respect to l persists; $\mu_B = e\hbar/(2m_q)$.

In this study *we ignore the spin* of the hydrogen's electron. The Zeeman effect *without the spin* of the electron is called the *normal* Zeeman effect. When the spin of the electron is considered, we get what is called the *anomalous* Zeeman effect, to be examined in Chapter 9 since its study requires familiarity with the formalisms of addition of angular momenta and perturbation theory, which will be studied in Chapters 7 and 9, respectively.

For simplicity, we take $\vec{B}$ along the z-direction: $\vec{B} = B\hat{z}$. The Hamiltonian of the hydrogen atom when subject to such a magnetic field can be obtained from (6.203) by replacing q with the electron's charge $q \to -e$,

$$\hat{H} = \frac{1}{2m_q}\hat{\vec{p}}^2 - \frac{e^2}{r} + \frac{e}{2m_q}B\hat{L}_z + \frac{e^2B^2}{8m_q}\left(x^2 + y^2\right) = \hat{H}_0 + \frac{e}{2m_q}B\hat{L}_z + \frac{e^2B^2}{8m_q}\left(x^2 + y^2\right), \tag{6.204}$$

where $\hat{H}_0 = \hat{\vec{p}}^2/(2m_q) - e^2/r$ is the atom's Hamiltonian in the absence of a magnetic field. We can ignore the quadratic term $e^2B^2(x^2 + y^2)/(8m_q)$; it is too small for a one-electron atom even when the field $\vec{B}$ is strong; then (6.204) reduces to

$$\hat{H} = \hat{H}_0 + \frac{B\mu_B}{\hbar}\hat{L}_z, \tag{6.205}$$

where $\mu_B = e\hbar/(2m_q) = 9.2740 \times 10^{-24}\,\mathrm{J\,T^{-1}} = 5.7884 \times 10^{-5}\,\mathrm{eV\,T^{-1}}$ is the Bohr magneton; the electron's orbital magnetic dipole moment, which results from the orbiting motion of the electron about the proton, would be given by $\vec{\mu}_L = -e\vec{B}/(2m_q)$. Since $\hat{H}_0$ commutes with $\hat{L}_z$, the operators $\hat{H}$, $\hat{L}_z$, and $\hat{H}_0$ mutually commute; hence they possess a set of common

eigenfunctions: $\psi_{nlm}(r,\theta,\varphi) = R_{nl}(r)Y_{lm}(\theta,\varphi)$. The eigenvalues of (6.205) are

$$E_{nlm} = \langle nlm \mid \hat{H} \mid nlm\rangle = \langle nlm \mid \hat{H}_0 \mid nlm\rangle + \frac{B\mu_B}{\hbar}\langle nlm \mid \hat{L}_z \mid nlm\rangle \tag{6.206}$$

or

$$E_{nlm} = E_n^0 + m\mu_B B = E_n^0 + m\hbar\omega_L, \tag{6.207}$$

where E_n^0 are the hydrogen's energy levels $E_n^0 = -m_q e^4/(2\hbar^2 n^2)$ (6.147) and ω_L is called the *Larmor frequency*:

$$\omega_L = \frac{eB}{2m_q} \equiv \frac{\mu_B}{\hbar}B. \tag{6.208}$$

So when a hydrogen atom is placed in a uniform magnetic field, and if we ignore the spin of the electron, the atom's spherical symmetry will be broken: each level with angular momentum l will split into $(2l+1)$ equally spaced levels (Figure 6.4), where the spacing is given by $\Delta E = \hbar\omega_L = B\mu_B$; the spacing is independent of l. This *equidistant* splitting of the levels is known as the *normal* Zeeman effect. The splitting leads to transitions which are restricted by the selection rule: $\Delta m = -1, 0, 1$. Transitions $m' = 0 \longrightarrow m = 0$ are not allowed.

The normal Zeeman effect has removed the degeneracy of the levels only partially; the degeneracy with respect to l remains. For instance, as shown in Figure 6.4, the following levels are still degenerate: $E_{nlm} = E_{200} = E_{210}$, $E_{32,-1} = E_{31,-1}$, $E_{300} = E_{310} = E_{320}$, and $E_{321} = E_{311}$. That is, the degeneracies of the levels corresponding to the same n and m but different values of l are not removed by the normal Zeeman effect: $E_{nl'm} = E_{nlm}$ with $l' \neq l$.

The results of the normal Zeeman effect, which show that each energy level splits into an *odd* number of $(2l+1)$ equally spaced levels, disagree with the experimental observations. For instance, every level in the hydrogen atom actually splits into an *even* number of levels. This suggests that the angular momentum is not integer but half-integer. This disagreement is due to the simplifying assumption where the spin of the electron was ignored. A proper treatment, which includes the electron spin, confirms that the angular momentum is not purely orbital but includes a spin component as well. This leads to the splitting of each level into an *even*[8] number of $(2j+1)$ unequally spaced energy levels. This effect, known as the *anomalous* Zeeman effect, is in full agreement with experimental findings.

6.4 Pauli Equation for a Spin 1/2 Particle in a Magnetic Field

It is important to note that the Schrödinger equation in its original form does not deal with spin; for instance, the Schrödinger Hamiltonian for a hydrogen atom, $\hat{H}_S = \hat{\vec{P}}^2/(2m) - e^2/r$, does not account for the spin degrees of freedom. Without accounting for spin, nonrelativistic quantum mechanics is incomplete. The project of expanding the nonrelativistic Schrödinger equation to include spin was carried out in great part by Pauli. Inspired by the 1922 Stern–Gerlach which confirmed experimentally the existence of the spin of the electron, Pauli attempted in 1927 to incorporate the interaction of the electron's spin with a magnetic field into the Schrödinger equation. He quickly encountered a number of conceptual as well as technical difficulties. Unable to find a relativistic version of the Schrödinger equation (this project had to wait for

[8] When spin is included, the electron's total angular momentum j would be half-integer; $(2j+1)$ is then an even number.

Dirac), he ended up settling for a less rigorous approach by adding spin to the Schrödinger equation in a rather ad hoc manner as we will see below; this work has given us a modified version of the Schrödinger equation called the Schrödinger–Pauli equation or, in short, the Pauli equation.

Guided by two discoveries – the 1922 Stern–Gerlach experiment that has confirmed that the bean of silver atoms passing through a non-uniform magnetic split into two distinct components and by the 1925 idea of Goudsmit and Uhlenbeck that this two-component splitting is due to the *spin 1/2* of the valence electron (in the silver atom) – Pauli has proposed in 1927 his *two component spinor theory* by invoking a set of three 2×2 matrices, the Pauli matrices that were discussed in Chapter 5 (see Section 5.6).

In this Subsection, we are going to derive the Pauli equation by proceeding in two steps: first, by considering a free spin 1/2 particle (with no external field) and, second, a charged spin 1/2 particle interacting with an external electromagnetic field.

6.4.1 Pauli equation for a free Spin 1/2 Particle

To account for spin 1/2 particles, Pauli introduced a set of three 2×2 matrices[9]:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.209}$$

In addition to being Hermitian and unitary, these matrices satisfy the following property:

$$\left\{\sigma_j, \sigma_k\right\} = 2\hat{I}_2\delta_{j,k}, \qquad (j, k = 1, 2, 3), \tag{6.210}$$

where $\hat{I}_2$ is the 2×2 unit matrix; hence, we see that $\sigma_j^2 = \hat{I}_2$ where $j = 1, 2, 3$. Using (6.210), we can verify that, for any two arbitrary vectors $\vec{A}$ and $\vec{B}$ which commute with $\vec{\sigma}$, we have

$$(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) = (\vec{A} \cdot \vec{B})\hat{I}_2 + i\vec{\sigma} \cdot (\vec{A} \times \vec{B}). \tag{6.211}$$

As seen in Chapter 5, the spin vector operator is given in terms of the Pauli matrices as follows:

$$\hat{\vec{S}} = \frac{\hbar}{2}\vec{\sigma}. \tag{6.212}$$

Clearly, both of $\hat{\vec{S}}$ and $\vec{\sigma}$ are 2×2 matrices.

Pauli equation for a free particle

Using (6.211) and since $\hat{\vec{p}} \times \hat{\vec{p}} = 0$, we have[10]

$$(\vec{\sigma} \cdot \hat{\vec{p}})(\vec{\sigma} \cdot \hat{\vec{p}}) = (\hat{\vec{p}} \cdot \hat{\vec{p}})\hat{I}_2 + i\vec{\sigma} \cdot (\hat{\vec{p}} \times \hat{\vec{p}}) \quad \Longrightarrow \quad (\vec{\sigma} \cdot \hat{\vec{p}})^2 = \hat{\vec{p}}^2\hat{I}_2 \equiv \hat{\vec{p}}^2. \tag{6.213}$$

This property has some historical significance: *it has enabled Pauli to introduce the spin into the Schrödinger equation without affecting its spatial degrees of freedom.* For instance, in the case of a free particle, the time independent Schrödinger equation

$$\hat{H} \mid \psi\rangle = \frac{\hat{\vec{p}}^2}{2m} \mid \psi\rangle, \tag{6.214}$$

[9] For notational brevity, we will use the Index Notation (see Appendix 4) where $\sigma_x \equiv \sigma_1$, $\sigma_y \equiv \sigma_2$, $\sigma_z \equiv \sigma_3$.

[10] A more explicit derivation of $(\vec{\sigma} \cdot \hat{\vec{p}})^2 = \hat{\vec{p}}^2$ is presented in Example 6.4.

can be "retooled" to describe a free spin 1/2 particle as follows:

$$\hat{H} \mid \phi\rangle = \frac{(\vec{\sigma}\cdot\hat{\vec{p}})^2}{2m} \mid \phi\rangle \equiv \frac{\hat{\vec{p}}^2}{2m} I_2 \mid \phi\rangle. \tag{6.215}$$

This is known as the *Pauli equation for a free spin 1/2 particle.* Since each one of the σ_j matrices is a 2×2 matrix, the state $\mid \phi\rangle$ must have two components $\mid \phi_A\rangle$ and $\mid \phi_B\rangle$ that represent the intrinsic spin degrees of freedom of the spin 1/2 particle:

$$\mid \phi\rangle = \begin{pmatrix} \mid \phi_A\rangle \\ \mid \phi_B\rangle \end{pmatrix}. \tag{6.216}$$

The state $\mid \phi\rangle$ is known as a *spinor*, or the *Pauli spinor.*

6.4.2 Pauli equation for a Spin 1/2 Particle in an Electromagnetic Field

Let us now consider a particle of charge q that is interacting with an electromagnetic field (having an electric scalar potential U and a vector potential $\vec{A}$). For this, we are going to invoke the *minimal coupling principle* principle also known as the *minimal substitution* principle; it is called "minimal" because it couples the electromagnetic field to the "pointlike" aspect of the particle. This principle states that the Schrödinger equation for a charged particle, q, that is interacting with an external electromagnetic field can be obtained by making the following substitutions[11]

$$\hat{H}_0 \longrightarrow \hat{H} = \hat{H}_0 - qU, \qquad \hat{\vec{p}}_0 \longrightarrow \hat{\vec{p}} = \hat{\vec{p}}_0 - q\vec{A}, \tag{6.217}$$

into the Schrödinger equation of the corresponding *free* particle, where $qU(\vec{r})$ is the electric energy due to the interaction between the charge and the electric scalar potential U; H_0 and p_0 are the Hamiltonian and momentum of the *free* particle, respectively. For simplicity, we are considering the case where there is no external potential $\hat{V}(\vec{r})$; hence, the charge interacts only with the external electromagnetic field. However, if, in addition to interacting with an external electromagnetic field, the particle is also subjected to an external potential $\hat{V}(\vec{r})$, the formalism below can handle it quite readily; all we need to do is to simply replace $qU(\vec{r})$ by $(qU(\vec{r}) + \hat{V}(\vec{r}))$ in all relevant expressions below.

If we insert the minimal substitution relations (6.217) into the free-particle Pauli equation (6.215), we obtain

$$\hat{H} \mid \phi\rangle = \frac{\left[\vec{\sigma}\cdot(\hat{\vec{p}} - q\vec{A})\right]^2}{2m} \mid \phi\rangle + qU \mid \phi\rangle. \tag{6.218}$$

Now, let us focus on the calculation of $\left[\vec{\sigma}\cdot\left(\hat{\vec{p}} - q\vec{A}\right)\right]^2$. Using the relation $\sigma_j\sigma_k = \delta_{jk}I_2 + i\varepsilon_{jkl}\sigma_l$ from Chapter 5, we can verify that, for any two arbitrary vectors $\vec{C}$ and $\vec{D}$ which commute with $\vec{\sigma}$, we have

$$(\vec{\sigma}\cdot\vec{C})(\vec{\sigma}\cdot\vec{D}) = (\vec{C}\cdot\vec{D})I_2 + i\vec{\sigma}\cdot(\vec{C}\times\vec{D}). \tag{6.219}$$

Using this relation, we can write

$$\left[\vec{\sigma}\cdot\left(\hat{\vec{p}} - q\vec{A}\right)\right]\left[\vec{\sigma}\cdot\left(\hat{\vec{p}} - q\vec{A}\right)\right] = \left(\hat{\vec{p}} - q\vec{A}\right)^2 + i\vec{\sigma}\cdot\left[\left(\hat{\vec{p}} - q\vec{A}\right)\times\left(\hat{\vec{p}} - q\vec{A}\right)\right], \tag{6.220}$$

[11] In Chapter 12, we show how to obtain these relations from the classical Lagrangian of a charged particle in an electromagnetic field: see Eqs. (13.136)–(13.137).

where

$$\begin{aligned}\left(\hat{\vec{p}}-q\vec{A}\right)\times\left(\hat{\vec{p}}-q\vec{A}\right) &= -q\left(\hat{\vec{p}}\times\vec{A}+\vec{A}\times\hat{\vec{p}}\right)+\hat{\vec{p}}\times\hat{\vec{p}}+q^2\vec{A}\times\vec{A}\\ &= -q\left(\hat{\vec{p}}\times\vec{A}+\vec{A}\times\hat{\vec{p}}\right),\end{aligned} \tag{6.221}$$

where we have used $\hat{\vec{p}}\times\hat{\vec{p}}=0$ and $\vec{A}\times\vec{A}=0$ in the last step. Now, since $\hat{\vec{p}}=-i\hbar\vec{\nabla}$ and using the fact that the components of $\hat{\vec{p}}$ and $\vec{A}$ do not commute, we have

$$\hat{\vec{p}}\times\vec{A}+\vec{A}\times\hat{\vec{p}}=-i\hbar\left(\vec{\nabla}\times\vec{A}+\vec{A}\times\vec{\nabla}\right). \tag{6.222}$$

Next, using the identity[12] $\vec{\nabla}\times(\vec{A}\phi)=\left(\vec{\nabla}\times\vec{A}\right)\phi-\vec{A}\times(\vec{\nabla}\phi)$, we can write

$$\begin{aligned}\vec{\nabla}\times\vec{A}\phi+\vec{A}\times\vec{\nabla}\phi &= \left(\vec{\nabla}\times\vec{A}\right)\phi-\vec{A}\times(\vec{\nabla}\phi)+\vec{A}\times(\vec{\nabla}\phi)\\ &= \left(\vec{\nabla}\times\vec{A}\right)\phi\\ &\equiv \vec{B}\phi,\end{aligned} \tag{6.223}$$

where, in the last step, we have used the relation $\vec{B}=\vec{\nabla}\times\vec{A}$ which relates the magnetic field $\vec{B}$ to the vector potential $\vec{A}$. Now, plugging (6.223) into (6.222), we obtain

$$\hat{\vec{p}}\times\vec{A}+\vec{A}\times\hat{\vec{p}}=-i\hbar\vec{B}, \tag{6.224}$$

which, in turn, when inserted into the last term of (6.221) leads to

$$\left(\hat{\vec{p}}-q\vec{A}\right)\times\left(\hat{\vec{p}}-q\vec{A}\right)=-q\left(\hat{\vec{p}}\times\vec{A}+\vec{A}\times\hat{\vec{p}}\right)=iq\hbar\left(\vec{\nabla}\times\vec{A}\right)=iq\hbar\vec{B}, \tag{6.225}$$

Finally, substituting (6.225) into (6.220), we obtain

$$\left[\vec{\sigma}\cdot\left(\hat{\vec{p}}-q\vec{A}\right)\right]\left[\vec{\sigma}\cdot\left(\hat{\vec{p}}-q\vec{A}\right)\right]=\left(\hat{\vec{p}}-q\vec{A}\right)^2-q\hbar\vec{\sigma}\cdot\vec{B}, \tag{6.226}$$

which, in turn, when inserted into (6.218), leads to

$$\left[\frac{1}{2m}\left(\hat{\vec{p}}-q\vec{A}\right)^2+\hat{V}-\frac{q\hbar}{2m}\vec{\sigma}\cdot\vec{B}\right]|\,\phi\rangle=E\,|\,\phi\rangle. \tag{6.227}$$

Since the spin operator is given by $\hat{\vec{S}}=(\hbar/2)\vec{\sigma}$, we can rewrite this equation as

$$\boxed{\left[\frac{1}{2m}\left(\hat{\vec{p}}-q\vec{A}\right)^2+qU-\frac{q}{m}\hat{\vec{S}}\cdot\vec{B}\right]|\,\phi\rangle=E\,|\,\phi\rangle.} \tag{6.228}$$

This is the *Pauli equation*[13] for a spin 1/2 particle of charge q in an electromagnetic field, where $|\,\phi\rangle$ is the (2–component) Pauli spinor; these two components correspond to the

[12] In this identity, the operator $\vec{\nabla}$ in the "$\left(\vec{\nabla}\times\vec{A}\right)\phi$" term acts only on $\vec{A}$, not on ϕ nor on anything else on the right-hand side.

[13] W. Pauli, *Z. Phys.* 43, 601 (1927).

two orientations of the particle's spin. This equation is also called the *Schrödinger–Pauli equation*. We will show in Chapter 12 that the Pauli equation is the non-relativistic limit of the Dirac equation. Note that if the magnetic vector potential $\vec{A}$ vanishes, the Pauli equation reduces to the familiar Schrödinger equation of a charged particle in an electric potential $U(\vec{r})$: $\left(\frac{1}{2m}\hat{\vec{p}}^2 + qU\right) \mid \phi\rangle = E \mid \phi\rangle$. The term $\frac{q}{m}\hat{\vec{S}} \cdot \vec{B}$, which results from the interaction between the electron's spin and the external magnetic field, is known as the Stern–Gerlach term.

The first term in (6.228) is the usual kinetic term of the Schrödinger equation. As we will see in Chapter 9 (see Subsection 9.2.3.1), the term $-(q/m)\hat{\vec{S}} \cdot \vec{B}$ represents the interaction energy of the particle's magnetic dipole moment $\vec{\mu}_S$ with an external magnetic field $\vec{B}$:

$$\hat{H}_S = -\frac{q}{m}\hat{\vec{S}} \cdot \vec{B} \equiv -\hat{\vec{\mu}}_S \cdot \vec{B} \quad \Longrightarrow \quad \hat{\vec{\mu}}_S = \frac{q}{m}\hat{\vec{S}}, \tag{6.229}$$

where $\hat{\vec{\mu}}_S$ is known as the intrinsic or *spin magnetic moment* of the spin 1/2 particle. This term is also known as the spin-orbit coupling as we will see in Chapter 9 because it is proportional to the scalar product between the spin and orbital angular momenta: $\hat{H}_S \propto \hat{\vec{S}} \cdot \hat{\vec{L}}$ where $\hat{\vec{L}}$ is the orbital angular momentum of the particle.

We can rewrite the Pauli equation (6.228) in a differential form as follows

$$\boxed{i\hbar \frac{\partial |\phi\rangle}{\partial t} = \left[\frac{1}{2m}\left(\hat{\vec{p}} - q\vec{A}\right)^2 + qU - \hat{\vec{\mu}}_S \cdot \vec{B}\right] \mid \phi\rangle.} \tag{6.230}$$

We may recall from Chapter 5 (see Subsection 5.6.1) that the classical orbital magnetic moment is given by

$$\hat{\vec{\mu}}_L = \frac{q}{2m}\hat{\vec{L}}. \tag{6.231}$$

A comparison of (6.229) and (6.231) shows that the spin magnetic moment of a spin 1/2 particle is *twice* that obtained from classical physics. As such, we can express $\hat{\vec{\mu}}_S$ in terms of a factor called the g–factor or the *gyromagnetic ratio* $g_s = 2$:

$$\boxed{\hat{\vec{\mu}}_S = \frac{q}{m}\hat{\vec{S}} = g_s \frac{q}{2m}\hat{\vec{S}} \quad \Longrightarrow \quad \text{gyromagnetic ratio} \quad g_s = 2.} \tag{6.232}$$

Although Pauli has added the spin degree of freedom to Schrödinger equation in an ad hoc manner, the Pauli equation has undeniably succeeded in giving us the correct $g_s = 2$ value for the gyromagnetic ratio; this result is in amazing agreement with the experimental value of $g_{EXP} = 2.0023193\cdots$. We will show in Chapter 12 that the Dirac equation also gives the correct $g_s = 2$ value; however, unlike the Pauli equation, the Dirac equation accounts for the spin from the very outset, with no need to graft it onto the theory.

Example 6.4

Using the properties of the Pauli matrices, verify that $(\vec{\sigma} \cdot \vec{p})^2 = \vec{p}^2$.

Solution

Since $\vec{\sigma} \cdot \vec{p} = \sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3$, we have

$$\begin{aligned}(\vec{\sigma} \cdot \vec{p})^2 &= (\sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3)(\sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3) \\ &= (\sigma_1^2 p_1^2 + \sigma_2^2 p_2^2 + \sigma_3^2 p_3^2) \\ &+ (\sigma_1\sigma_2 p_1 p_2 + \sigma_2\sigma_1 p_2 p_1) + (\sigma_1\sigma_3 p_1 p_3 + \sigma_3\sigma_1 p_3 p_1) + (\sigma_2\sigma_3 p_2 p_3 + \sigma_3\sigma_2 p_3 p_2) \\ &= p_1^2 + p_2^2 + p_3^2 + (\sigma_1\sigma_2 + \sigma_2\sigma_1)p_1 p_2 + (\sigma_1\sigma_3 + \sigma_3\sigma_2)p_1 p_3 + (\sigma_2\sigma_3 + \sigma_3\sigma_2)p_2 p_3. \end{aligned} \tag{6.233}$$

where, in the last line, we have used the fact that $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \hat{I}_2$ and that the components of the momentum operator commute (i.e., $p_j p_k = p_k p_j$ for $j, k = 1, 2, 3$). Additionally, since the different components of the Pauli matrices anticommute (i.e., $\sigma_j \sigma_k = -\sigma_k \sigma_j$ for $j \neq k$), each of the last three terms of (6.233) is zero. Hence, we can reduce (6.233) to

$$(\vec{\sigma} \cdot \vec{p})^2 = p_1^2 + p_2^2 + p_3^2. \tag{6.234}$$

6.5 Concluding Remarks

An important result that needs to be highlighted in this chapter is the solution of the Schrödinger equation for the hydrogen atom and how it has solved the conceptual as well as the quantitative issues that have plagued the old quantum theory during the first quarter of the 20th century; that momentous era was dominated in great part by the Bohr model of the atom. To appreciate the power of the Schrödinger equation, let us take a brief look at its success in addressing some of the deficiencies of the Bohr model.

- Unlike Bohr's semiclassical model[14] which is founded on piecemeal assumptions, we have seen in this chapter how the Schrödinger equation offers a mathematically rigorous description of the electrons; it yields the quantized expressions for the atom's radius and energy levels without resorting to any ad hoc argument. Once the Schrödinger equation was established, we have allowed the mathematical machinery of differential equations to run its logical course; no hand-waving steps were involved. For instance, the quantization of the energy levels comes out naturally as a by-product of the formalism, not resulting from an unjustified assumption: it is a consequence of the boundary conditions which require the wave function to be finite as $r \to \infty$; see equations (6.144) and (6.147).

- In its original form, the Bohr model is one-dimensional, for it uses only one quantum number to describe the distribution of electrons in an atom: only the principal quantum number "n" is needed to describe the size (radius) of the electron's orbit. In sharp contrast, Schrodinger's wave mechanics requires *three* quantum numbers (n, l, m) to describe the distribution of electrons in an atom because the state of an atomic electron

[14] As we saw in Chapter 1 (see Subsection 1.6.2), to explain the stability of atoms, Bohr introduced three rather arbitrary postulates that violated classical physics; after mixing them with a number of semiclassical ideas, Bohr proposed his *"quantized"* shell model for the hydrogen atom. Positing (without providing any proof) that the electron's orbital angular momentum is quantized, $L = n\hbar$, and using Coulomb's law and Newton's second law (i.e., $F = e^2/r^2 = mv^2/r$), Bohr managed to obtain *quantized expressions* for the atom's radius and energy levels.

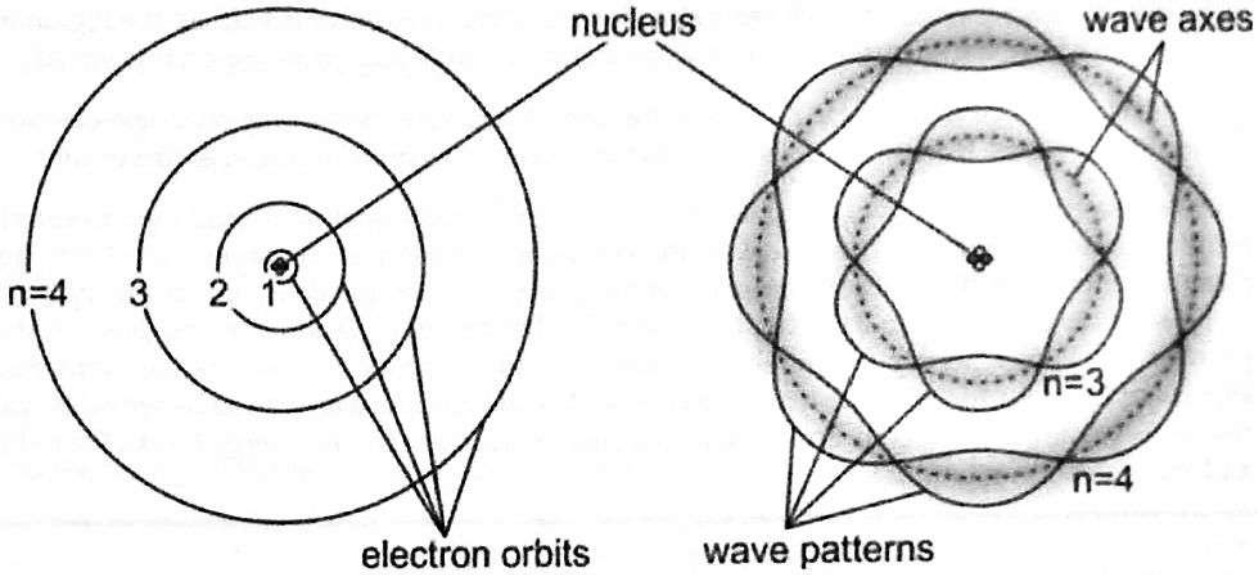


Figure 6.5: Bohr model versus the quantum mechanical model of the atom: **(Left)** the Bohr model treats electrons as point-like particles that rotate around the nucleus on "sharp" circular orbits, and **(Right)** the Schrödinger model treats electrons as three-dimensional standing waves that spread around the nucleus like clouds (electron clouds).

is described by means of a wave function, $\psi_{n,l,m}(\vec{r})$, that depends on three quantum numbers (n, l, m); this is expected since three coordinates (or, equivalently, three quantum numbers) are required to describe an electron in a three-dimensional space.

- The description of the electron by the Bohr model is conceptually flawed. As shown in Figure (6.5)-(Left), the electron is treated as a *point-like* particle that rotates around the nucleus on *"sharp"*, concentric circular orbits and that the *positions* of the electron are *well defined*. In the quantum mechanical model, however, since the electron is considered to display both particle and wave-like features as per the de Broglie matter-wave hypothesis, the electron is neither a point-like particle nor does it travel on a "sharp" path around the nucleus. In the Schrödinger wave mechanics view, the electron is represented by a three-dimensional *standing wave* that must have an *integer* number of wavelengths as illustrated in Figure (6.5)-(Right). The three-dimensional standing waves spread around the nucleus like *clouds*[15] (see the shaded areas in Figure 6.5-(Right)); the shapes of the clouds are delimited by the standing waves that are formed by the electrons trapped inside them. In sharp contrast to the *deterministic* nature of the Bohr model, only a *probabilistic* outcome is possible within the framework of quantum mechanics: we cannot know exactly where an electron is at any given time, but we know the region where it is likely to be found. The probability density — i.e., the square modulus of the electron's wave function, $\left|\psi_{n,l,m}(\vec{r})\right|^2$ — of finding an electron is highest in the region of space where the electron cloud is densest (i.e., at the crests or antinodes of the standing wave) and smallest where the cloud is thinnest (i.e., at the nodes). So

[15]Since the electrons travel so fast (at speeds approaching the speed of light) on spherical shells along all kinds of angles around the nucleus, they can be viewed as forming clouds around the nucleus. The *electron cloud model* provides a practical visual illustration on how electrons travel around the nucleus.

electrons are most likely to be found inside the dense regions of the cloud; the region in space where the electron can be most probably found is known as the *electron orbital*.

- While the Bohr model calculations of the energy and radii work only on single-electron atoms or ions, the Schrödinger equation applies also to multi-electron atoms or ions.

From a historical perspective, the primary motivator that has impelled Schrödinger to search for a consistent theory that can explain the behavior of matter at the atomic level was his dissatisfaction with the arbitrary nature and the guess-work of the old quantum theory.

In summary, we saw in this chapter how the Schrödinger equation has delivered on the promise made in Chapter 1; namely, to come up with a theory that avoids the undesired aspects of the Bohr model —its hand-waving, ad hoc assumptions as well as its semiclassical nature — while preserving its good points (i.e., the expressions for the energy levels, the radii, and the transition relations).

6.6 Solved Problems

Problem 6.1
Consider a spinless particle of mass m which is moving in a three-dimensional potential

$$V(x,y,z) = \begin{cases} \frac{1}{2}m\omega^2 z^2, & 0 < x < a,\ 0 < y < a, \\ \infty, & \text{elsewhere.} \end{cases}$$

(a) Write down the total energy and the total wave function of this particle.

(b) Assuming that $\hbar\omega > 3\pi^2\hbar^2/(2ma^2)$, find the energies and the corresponding degeneracies for the ground state and first excited state.

(c) Assume now that, in addition to the potential $V(x,y,z)$, this particle also has a negative electric charge $-q$ and that it is subjected to a constant electric field ϵ directed along the z-axis. The Hamiltonian along the z-axis is thus given by

$$\hat{H}_z = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + \frac{1}{2}m\omega^2 z^2 - q\epsilon z.$$

Derive the energy expression E_{n_z} for this particle and also its total energy $E_{n_x n_y n_z}$. Then find the energies and the corresponding degeneracies for the ground state and first excited state.

Solution

(a) This three-dimensional potential consists of three independent one-dimensional potentials: (i) a potential well along the x-axis, (ii) a potential well along the y-axis, and (iii) a harmonic oscillator along the z-axis. The energy must then be given by

$$E_{n_x n_y n_z} = \frac{\pi^2\hbar^2}{2ma^2}\left(n_x^2 + n_y^2\right) + \hbar\omega\left(n_z + \frac{1}{2}\right) \tag{6.235}$$

and the wave function by

$$\psi_{n_x n_y n_z}(x,y,z) = X_{n_x}(x)Y_{n_y}(y)Z_{n_z}(z) = \frac{2}{a}\sin\left(\frac{\pi n_x}{a}x\right)\sin\left(\frac{\pi n_y}{a}y\right)Z_{n_z}, \tag{6.236}$$

where $Z_{n_z}(z)$ is the wave function of a harmonic oscillator which, as shown in Chapter 4, is given in terms of the Hermite polynomial $H_{n_z}\left(\frac{z}{z_0}\right)$ by

$$Z_{n_z}(z) = \frac{1}{\sqrt{\sqrt{\pi}2^{n_z}n_z!z_0}} e^{-z^2/2z_0^2} H_{n_z}\left(\frac{z}{z_0}\right), \tag{6.237}$$

with $z_0 = \sqrt{\hbar/(m\omega)}$.

(b) The energy of the ground state is given by

$$E_{110} = \frac{\pi^2\hbar^2}{ma^2} + \frac{\hbar\omega}{2} \tag{6.238}$$

and the energy of the first excited state is given by

$$E_{120} = E_{210} = \frac{5\pi^2\hbar^2}{2ma^2} + \frac{\hbar\omega}{2}. \tag{6.239}$$

Note that, while the ground state is not degenerate, the first excited state is twofold degenerate. We should also mention that, since $\hbar\omega > 3\pi^2\hbar^2/(2ma^2)$, we have $E_{120} < E_{111}$, or

$$E_{111} = \frac{\pi^2\hbar^2}{ma^2} + \frac{3\hbar\omega}{2} = E_{120} + \hbar\omega - \frac{3\pi^2\hbar^2}{2ma^2}, \tag{6.240}$$

and hence the first excited state is given by E_{120} and not by E_{111}.

(c) To obtain the energies for

$$\hat{H}_z = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + \frac{1}{2}m\omega^2 z^2 - q\epsilon z, \tag{6.241}$$

we need simply to make the change of variable $\lambda = z - q\epsilon/(m\omega^2)$; hence $dz = d\lambda$. The Hamiltonian $\hat{H}_z$ then reduces to

$$\hat{H}_z = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial \lambda^2} + \frac{1}{2}m\omega^2\lambda^2 - \frac{q^2\epsilon^2}{2m\omega^2}. \tag{6.242}$$

This suggestive form implies that the energy eigenvalues of $\hat{H}_z$ are those of a harmonic oscillator that are shifted downwards by an amount equal to $q^2\epsilon^2/(2m\omega^2)$:

$$E_{n_z} = \langle n_z|\hat{H}_z|n_z\rangle = \hbar\omega\left(n_z + \frac{1}{2}\right) - \frac{q^2\epsilon^2}{2m\omega^2}. \tag{6.243}$$

As a result, the total energy is now given by

$$E_{n_x n_y n_z} = \frac{\pi^2\hbar^2}{2ma^2}\left(n_x^2 + n_y^2\right) + \hbar\omega\left(n_z + \frac{1}{2}\right) - \frac{q^2\epsilon^2}{2m\omega^2}. \tag{6.244}$$

The energies of the ground and first excited states are

$$E_{110} = \frac{\pi^2\hbar^2}{ma^2} + \frac{\hbar\omega}{2} - \frac{q^2\epsilon^2}{2m\omega^2}, \qquad E_{120} = E_{210} = \frac{5\pi^2\hbar^2}{2ma^2} + \frac{\hbar\omega}{2} - \frac{q^2\epsilon^2}{2m\omega^2}. \tag{6.245}$$

Problem 6.2
Show how to obtain the expressions of: (a) $\langle nl|r^{-2}|nl\rangle$ and (b) $\langle nl|r^{-1}|nl\rangle$; that is, prove (6.183) and (6.182).

Solution
The starting point is the radial equation (6.127),

$$-\frac{\hbar^2}{2\mu}\frac{d^2U_{nl}(r)}{dr^2}+\left[\frac{l(l+1)\hbar^2}{2\mu r^2}-\frac{e^2}{r}\right]U_{nl}(r)=E_nU_{nl}(r), \tag{6.246}$$

which can be rewritten as

$$\frac{U''_{nl}(r)}{U_{nl}(r)}=\frac{l(l+1)}{r^2}-\frac{2\mu e^2}{\hbar^2}\frac{1}{r}+\frac{\mu^2e^4}{\hbar^4n^2}, \tag{6.247}$$

where $U_{nl}(r)=rR_{nl}(r)$, $U''_{nl}(r)=d^2U_{nl}(r)/dr^2$, and $E_n=-\mu e^4/(2\hbar^2n^2)$.

(a) To find $\left\langle r^{-2}\right\rangle_{nl}$, let us treat the orbital quantum number l as a continuous variable and take the first l derivative of (6.247):

$$\frac{\partial}{\partial l}\left[\frac{U''_{nl}(r)}{U_{nl}(r)}\right]=\frac{2l+1}{r^2}-\frac{2\mu^2e^4}{\hbar^4n^3}, \tag{6.248}$$

where we have the fact that n depends on l since, as shown in (6.145), $n=N+l+1$; thus $\partial n/\partial l=1$. Now since $\int_0^\infty U_{nl}^2(r)\,dr=\int_0^\infty r^2R_{nl}^2(r)\,dr=1$, multiplying both sides of (6.248) by $U_{nl}^2(r)$ and integrating over r we get

$$\int_0^\infty U_{nl}^2(r)\frac{\partial}{\partial l}\left[\frac{U''_{nl}(r)}{U_{nl}(r)}\right]dr=(2l+1)\int_0^\infty U_{nl}^2(r)\frac{1}{r^2}dr-\frac{2\mu^2e^4}{\hbar^4n^3}\int_0^\infty U_{nl}^2(r)\,dr, \tag{6.249}$$

or

$$\int_0^\infty U_{nl}^2(r)\frac{\partial}{\partial l}\left[\frac{U''_{nl}(r)}{U_{nl}(r)}\right]dr=(2l+1)\left\langle nl\left|\frac{1}{r^2}\right|nl\right\rangle-\frac{2\mu^2e^4}{\hbar^4n^3}. \tag{6.250}$$

The left-hand side of this relation is equal to zero, since

$$\int_0^\infty U_{nl}^2(r)\frac{\partial}{\partial l}\left[\frac{U''_{nl}(r)}{U_{nl}(r)}\right]dr=\int_0^\infty U_{nl}(r)\frac{\partial U''_{nl}(r)}{\partial l}dr-\int_0^\infty U''_{nl}(r)\frac{\partial U_{nl}(r)}{\partial l}dr=0. \tag{6.251}$$

We may therefore rewrite (6.250) as

$$(2l+1)\left\langle nl\left|\frac{1}{r^2}\right|nl\right\rangle=\frac{2\mu^2e^4}{\hbar^4n^3}; \tag{6.252}$$

hence

$$\left\langle nl\left|\frac{1}{r^2}\right|nl\right\rangle=\frac{2}{n^3(2l+1)a_0^2}, \tag{6.253}$$

since $a_0=\hbar^2/(\mu e^2)$.

(b) To find $\langle r^{-1}\rangle_{nl}$ we need now to treat the electron's charge e as a continuous variable in (6.247). The first e-derivative of (6.247) yields

$$\frac{\partial}{\partial e}\left[\frac{U''_{nl}(r)}{U_{nl}(r)}\right] = -\frac{4\mu e}{\hbar^2}\frac{1}{r} + \frac{4\mu^2 e^3}{\hbar^4 n^2}. \tag{6.254}$$

Again, since $\int_0^\infty U_{nl}^2(r)\,dr = 1$, multiplying both sides of (6.254) by $U_{nl}^2(r)$ and integrating over r we obtain

$$\int_0^\infty U_{nl}^2(r)\frac{\partial}{\partial e}\left[\frac{U''_{nl}(r)}{U_{nl}(r)}\right]dr = -\frac{4\mu e}{\hbar^2}\int_0^\infty U_{nl}^2(r)\frac{1}{r}dr + \frac{4\mu^2 e^3}{\hbar^4 n^2}\int_0^\infty U_{nl}^2(r)\,dr, \tag{6.255}$$

or

$$\int_0^\infty U_{nl}^2(r)\frac{\partial}{\partial e}\left[\frac{U''_{nl}(r)}{U_{nl}(r)}\right]dr = -\frac{4\mu e}{\hbar^2}\left\langle nl\left|\frac{1}{r}\right|nl\right\rangle + \frac{4\mu^2 e^3}{\hbar^4 n^2}. \tag{6.256}$$

As shown in (6.251), the left-hand side of this is equal to zero. Thus, we have

$$\frac{4\mu e}{\hbar^2}\left\langle nl\left|\frac{1}{r}\right|nl\right\rangle = \frac{4\mu^2 e^3}{\hbar^4 n^2} \quad\Longrightarrow\quad \left\langle nl\left|\frac{1}{r}\right|nl\right\rangle = \frac{1}{n^2 a_0}, \tag{6.257}$$

since $a_0 = \hbar^2/(\mu e^2)$.

Problem 6.3

(a) Use Kramers' recursion rule (6.184) to obtain expressions (6.180) to (6.182) for $\langle nl|r^{-1}|nl\rangle$, $\langle nl|r|nl\rangle$, and $\langle nl|r^2|nl\rangle$.

(b) Using (6.253) for $\langle nl|r^{-2}|nl\rangle$ and combining it with Kramers' rule, obtain the expression for $\langle nl|r^{-3}|nl\rangle$.

(c) Repeat (b) to obtain the expression for $\langle nl|r^{-4}|nl\rangle$.

Solution

(a) First, to obtain $\left\langle nl\left|r^{-1}\right|nl\right\rangle$, we need simply to insert $k = 0$ into Kramers' recursion rule (6.184):

$$\frac{1}{n^2}\left\langle nl\left|r^0\right|nl\right\rangle - a_0\left\langle nl\left|r^{-1}\right|nl\right\rangle = 0; \tag{6.258}$$

hence

$$\left\langle nl\left|\frac{1}{r}\right|nl\right\rangle = \frac{1}{n^2 a_0}. \tag{6.259}$$

Second, an insertion of $k = 1$ into (6.184) leads to the relation for $\langle nl\,|r|\,nl\rangle$:

$$\frac{2}{n^2}\langle nl\,|r|\,nl\rangle - 3a_0\left\langle nl\left|r^0\right|nl\right\rangle + \frac{a_0^2}{4}\left[(2l+1)^2 - 1\right]\left\langle nl\left|r^{-1}\right|nl\right\rangle = 0, \tag{6.260}$$

and since $\left\langle nl\left|r^{-1}\right|nl\right\rangle = 1/(n^2 a_0)$, we have

$$\langle nl\,|r|\,nl\rangle = \frac{1}{2}\left[3n^2 - l(l+1)\right]a_0. \tag{6.261}$$

Third, substituting $k = 2$ into (6.184) we get

$$\frac{3}{n^2}\left\langle nl\left|r^2\right|nl\right\rangle - 5a_0\left\langle nl\left|r\right|nl\right\rangle + \frac{a_0^2}{2}\left[(2l+1)^2-4\right]\left\langle nl\left|r^0\right|nl\right\rangle = 0, \tag{6.262}$$

which when combined with $\left\langle nl\left|r\right|nl\right\rangle = \frac{1}{2}\left[3n^2 - l(l+1)\right]a_0$ yields

$$\boxed{\left\langle nl\left|r^2\right|nl\right\rangle = \frac{1}{2}n^2\left[5n^2 + 1 - 3l(l+1)\right]a_0^2.} \tag{6.263}$$

We can continue in this way to obtain any positive power of r: $\left\langle nl\left|r^k\right|nl\right\rangle$.

(b) Inserting $k = -1$ into Kramers' rule,

$$0 + a_0\left\langle nl\left|r^{-2}\right|nl\right\rangle - \frac{1}{4}\left[(2l+1)^2-1\right]a_0^2\left\langle nl\left|r^{-3}\right|nl\right\rangle, \tag{6.264}$$

we obtain

$$\left\langle nl\left|\frac{1}{r^3}\right|nl\right\rangle = \frac{1}{l(l+1)a_0}\left\langle nl\left|\frac{1}{r^2}\right|nl\right\rangle, \tag{6.265}$$

where the expression for $\left\langle nl\left|r^{-2}\right|nl\right\rangle$ is given by (6.253); thus, we have

$$\boxed{\left\langle nl\left|\frac{1}{r^3}\right|nl\right\rangle = \frac{2}{n^3l(l+1)(2l+1)a_0^3}.} \tag{6.266}$$

(c) To obtain the expression for $\left\langle nl\left|r^{-4}\right|nl\right\rangle$ we need to substitute $k = -2$ into Kramers' rule:

$$-\frac{1}{n^2}\left\langle nl\left|r^{-2}\right|nl\right\rangle + 3a_0\left\langle nl\left|r^{-3}\right|nl\right\rangle - \frac{a_0^2}{2}\left[(2l+1)^2-4\right]\left\langle nl\left|r^{-4}\right|nl\right\rangle = 0. \tag{6.267}$$

Inserting (6.253) and (6.266) for $\left\langle nl\left|r^{-2}\right|nl\right\rangle$ and $\left\langle nl\left|r^{-3}\right|nl\right\rangle$, we obtain

$$\boxed{\left\langle nl\left|\frac{1}{r^4}\right|nl\right\rangle = \frac{4\left[3n^2 - l(l+1)\right]}{n^5l(l+1)(2l+1)\left[(2l+1)^2-4\right]a_0^4}.} \tag{6.268}$$

We can continue in this way to obtain any negative power of r: $\left\langle nl\left|r^{-k}\right|nl\right\rangle$.

Problem 6.4

An electron is trapped inside an infinite spherical well $V(r) = \begin{cases} 0, & r < a, \\ +\infty, & r > a. \end{cases}$

(a) Using the radial Schrödinger equation, determine the bound eigenenergies and the corresponding normalized radial wave functions for the case where the orbital angular momentum of the electron is zero (i.e., $l = 0$).

(b) Show that the lowest energy state for $l = 7$ lies above the second lowest energy state for $l = 0$.

(c) Calculate the probability of finding the electron in a sphere of radius $a/2$, and then in a spherical shell of thickness $a/2$ situated between $r = a$ and $r = 3a/2$.

Solution

(a) Since $V(r) = 0$ in the region $r \leq a$, the radial Schrödinger equation (6.57) becomes

$$-\frac{\hbar^2}{2m}\left[\frac{d^2U_{nl}(r)}{dr^2} - \frac{l(l+1)}{r^2}U_{nl}(r)\right] = EU_{nl}(r), \tag{6.269}$$

where $U_{nl}(r) = rR_{nl}(r)$. For the case where $l = 0$, this equation reduces to

$$\frac{d^2U_{n0}(r)}{dr^2} = -k_n^2 U_{n0}(r), \tag{6.270}$$

where $k_n^2 = 2mE_n/\hbar^2$. The general solution to this differential equation is given by

$$U_{n0}(r) = A\cos(k_n r) + B\sin(k_n r) \tag{6.271}$$

or

$$R_{n0}(r) = \frac{1}{r}\left(A\cos(k_n r) + B\sin(k_n r)\right). \tag{6.272}$$

Since $R_{n0}(r)$ is finite at the origin or $U_{n0}(0) = 0$, the coefficient A must be zero. In addition, since the potential is infinite at $r = a$ (rigid wall), the radial function $R_{n0}(a)$ must vanish:

$$R_{n0}(a) = B\frac{\sin k_n a}{a} = 0; \tag{6.273}$$

hence $ka = n\pi$, $n = 1, 2, 3, \ldots$. This relation leads to

$$E_n = \frac{\hbar^2\pi^2}{2ma^2}n^2. \tag{6.274}$$

The normalization of the radial wave function $R(r)$, $\int_0^a |R_{n0}(r)|^2 r^2 dr = 1$, leads to

$$\begin{aligned} 1 &= |B|^2 \int_0^a \frac{1}{r^2}\sin^2(k_n r) r^2 dr = \frac{|B|^2}{k_n}\int_0^{k_n a} \sin^2\rho\, d\rho = \frac{|B|^2}{k_n}\left(\frac{\rho}{2} - \frac{\sin 2\rho}{4}\right)\Bigg|_{\rho=0}^{\rho=k_n a} \\ &= \frac{1}{2}|B|^2 a; \end{aligned} \tag{6.275}$$

hence $B = \sqrt{2/a}$. The normalized radial wave function is thus given by

$$R_{n0}(r) = \sqrt{\frac{2}{a}}\frac{1}{r}\sin\left(\sqrt{\frac{2mE_n}{\hbar^2}}r\right). \tag{6.276}$$

(b) For $l = 7$ we have

$$E_1(l = 7) > V_{eff}(l = 7) = \frac{56\hbar^2}{2ma^2} = \frac{28\hbar^2}{ma^2}. \tag{6.277}$$

The second lowest state for $l = 0$ is given by the 3s state; its energy is

$$E_2(l = 0) = \frac{2\pi^2\hbar^2}{ma^2}, \tag{6.278}$$

since $n = 2$. We see that

$$E_1(l = 7) > E_2(l = 0). \tag{6.279}$$

(c) Since the probability of finding the electron in the sphere of radius a is equal to 1, the probability of finding it in a sphere of radius $a/2$ is equal to 1/2.

As for the probability of finding the electron in the spherical shell between $r = a$ and $r = 3a/2$, it is equal to zero, since the electron cannot tunnel through the infinite potential from $r < a$ to $r > a$.

Problem 6.5

Find the $l = 0$ energy and wave function of a particle of mass m that is subject to the following central potential $V(r) = \begin{cases} 0, & a < r < b, \\ \infty, & \text{elsewhere.} \end{cases}$

Solution

This particle moves between two concentric, hard spheres of radii $r = a$ and $r = b$. The $l = 0$ radial equation between $a < r < b$ can be obtained from (6.57):

$$\frac{d^2U_{n0}(r)}{dr^2} + k^2U_{n0}(r) = 0, \tag{6.280}$$

where $U_{n0}(r) = rR_{n0}(r)$ and $k^2 = 2mE/\hbar^2$. Since the solutions of this equation must satisfy the condition $U_{n0}(a) = 0$, we may write

$$U_{n0}(r) = A\sin[k(r - a)]; \tag{6.281}$$

the radial wave function is zero elsewhere, i.e., $U_{n0}(r) = 0$ for $0 < r < a$ and $r > b$.

Moreover, since the radial function must vanish at $r = b$, $U_{n0}(b) = 0$, we have

$$A\sin[k(b - a)] = 0 \quad \Longrightarrow \quad k(b - a) = n\pi, \qquad n = 1, 2, 3, \ldots. \tag{6.282}$$

Coupled with the fact that $k^2 = 2mE/(\hbar^2)$, this condition leads to the energy

$$E_n = \frac{\hbar^2k^2}{2m} = \frac{n^2\pi^2\hbar^2}{2m(a - b)^2}, \qquad n = 1, 2, 3, \ldots. \tag{6.283}$$

We can normalize the radial function (6.281) to obtain the constant A:

$$\begin{aligned} 1 &= \int_a^b r^2R_{n0}^2(r)dr = \int_a^b U_{n0}^2(r)dr = A^2\int_a^b \sin^2[k(r - a)]\,dr \\ &= \frac{A^2}{2}\int_a^b \{1 - \cos[2k(r - a)]\}\,dr = \frac{b - a}{2}A^2; \end{aligned} \tag{6.284}$$

hence $A = \sqrt{2/(b - a)}$. Since $k_n = n\pi/(b - a)$ the normalized radial function is given by

$$R_{n0}(r) = \frac{1}{r}U_{n0}(r) = \begin{cases} \sqrt{\frac{2}{b-a}}\frac{1}{r}\sin[\frac{n\pi(r-a)}{b-a}], & a < r < b, \\ 0, & \text{elsewhere.} \end{cases} \tag{6.285}$$

To obtain the total wave function $\psi_{nlm}(\vec{r})$, we need simply to divide the radial function by a $1/\sqrt{4\pi}$ factor, because in this case of $l = 0$ the wave function $\psi_{n00}(r)$ depends on no angular degrees of freedom, it depends only on the radius:

$$\psi_{n00}(r) = \frac{1}{\sqrt{4\pi}} R_{n0}(r) = \begin{cases} \sqrt{\frac{2}{4\pi(b-a)}} \frac{1}{r} \sin[\frac{n\pi(r-a)}{b-a}], & a < r < b, \\ 0, & \text{elsewhere.} \end{cases} \tag{6.286}$$

Problem 6.6

(a) For the following cases, calculate the value of r at which the radial probability density of the hydrogen atom reaches its maximum: (i) $n = 1, l = 0, m = 0$; (ii) $n = 2, l = 1, m = 0$; (iii) $l = n - 1, m = 0$.

(b) Compare the values obtained with the Bohr radius for circular orbits.

Solution

(a) Since the radial wave function for $n = 1$ and $l = 0$ is $R_{10}(r) = 2a_0^{-3/2} e^{-r/a_0}$, the probability density is given by

$$P_{10}(r) = r^2 |R_{10}(r)|^2 = \frac{4}{a_0^3} r^2 e^{-2r/a_0}. \tag{6.287}$$

(i) The maximum of $P_{10}(r)$ occurs at r_1:

$$\left.\frac{dP_{10}(r)}{dr}\right|_{r=r_1} = 0 \implies 2r_1 - \frac{2r_1^2}{a_0} = 0 \implies r_1 = a_0. \tag{6.288}$$

(ii) Similarly, since $R_{21}(r) = 1/(2\sqrt{6}a_0^{5/2}) r e^{-r/2a_0}$, we have

$$P_{21}(r) = r^2 |R_{21}(r)|^2 = \frac{1}{24a_0^5} r^4 e^{-r/a_0}. \tag{6.289}$$

The maximum of the probability density is given by

$$\left.\frac{dP_{21}(r)}{dr}\right|_{r=r_2} = 0 \implies 4r_2^3 - \frac{r_2^4}{a_0} = 0 \implies r_2 = 4a_0. \tag{6.290}$$

(iii) The radial function for $l = n - 1$ can be obtained from (6.170):

$$R_{n(n-1)}(r) = -\left(\frac{2}{na_0}\right)^{3/2} \frac{1}{\sqrt{2n[(2n-1)!]^3}} \left(\frac{2r}{na_0}\right)^{(n-1)} e^{-r/na_0} L_{2n-1}^{2n-1}\left(\frac{2r}{na_0}\right). \tag{6.291}$$

From (6.159) and (6.160) we can verify that the associated Laguerre polynomial L_{2n-1}^{2n-1} is a *constant*, $L_{2n-1}^{2n-1}(y) = -(2n-1)!$. We can thus write $R_{n(n-1)}(r)$ as $R_{n(n-1)}(r) = A_n r^{n-1} e^{-r/na_0}$, where A_n is a constant. Hence the probability density is given by

$$P_{n(n-1)}(r) = r^2 |R_{n(n-1)}(r)|^2 = A_n^2 r^{2n} e^{-2r/na_0}. \tag{6.292}$$

The maximum of the probability density is given by

$$\left.\frac{dP_{n(n-1)}(r)}{dr}\right|_{r=r_n} = 0 \implies 2nr_n^{2n-1} - \frac{2r_n^{2n}}{na_0} = 0 \implies r_n = n^2 a_0. \tag{6.293}$$

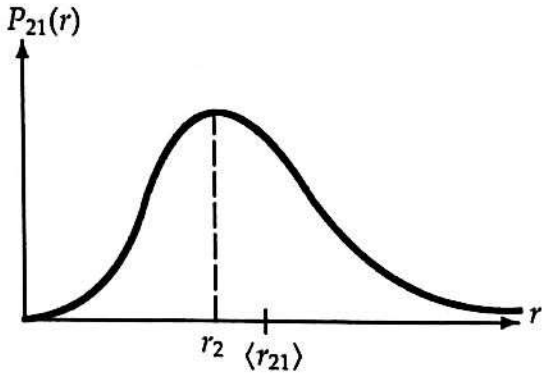


Figure 6.6: The probability density $P_{21}(r) = r^4 e^{-r/a_0}/(24a_0^5)$ is asymmetric about its maximum $r_2 = 4a_0$; the average of r is $\langle r_{21}\rangle = 5a_0$ and the width of the probability density is $\Delta r_{21} = \sqrt{5}a_0$.

(b) The values of r_n displayed in (6.288), (6.290), and (6.293) are nothing but the Bohr radii for circular orbits, $r_n = n^2 a_0$. The Bohr radius $r_n = n^2 a_0$ gives the position of maximum probability density for an electron in a hydrogen atom.

Problem 6.7

(a) Calculate the expectation value $\langle r\rangle_{21}$ for the hydrogen atom and compare it with the value r at which the radial probability density reaches its maximum for the state $n = 2$, $l = 1$.

(b) Calculate the width of the probability density distribution for r.

Solution

(a) Since $R_{21}(r) = re^{-r/2a_0}/\sqrt{24a_0^5}$ the average value of r in the state $R_{21}(r)$ is

$$\langle r\rangle_{21} = \frac{1}{24a_0^5}\int_0^\infty r^5 e^{-r/a_0}dr = \frac{a_0}{24}\int_0^\infty u^5 e^{-u}du = \frac{120a_0}{24} = 5a_0; \tag{6.294}$$

in deriving this relation we have made use of $\int_0^\infty x^n e^{-x}dx = n!$.

The value r at which the radial probability density reaches its maximum for the state $n = 2$, $l = 1$ is given by $r_2 = 4a_0$, as shown in (6.290).

What makes the results $r_2 = 4a_0$ and $\langle r\rangle_{21} = 5a_0$ different? The reason that $\langle r\rangle_{21}$ is different from r_2 can be attributed to the fact that the probability density $P_{21}(r)$ is asymmetric about its maximum, as shown in Figure 6.6. Although the most likely location of the electron is at $r_0 = 4a_0$, the average value of the measurement of its location is $\langle r\rangle_{21} = 5a_0$.

(b) The width of the probability distribution is given by $\Delta r = \sqrt{\langle r^2\rangle_{21} - \langle r\rangle_{21}^2}$, where the expectation value of r^2 is

$$\langle r^2\rangle_{21} = \int_0^\infty r^4 R_{21}^2(r)dr = \frac{1}{24a_0^5}\int_0^\infty r^6 \exp\left(-\frac{1}{a_0}r\right)dr = \frac{6!a_0^7}{24a_0^5} = 30a_0^2. \tag{6.295}$$

Thus, the width of the probability distribution shown in Figure 6.6 is given by

$$\Delta r_{21} = \sqrt{\langle r^2\rangle_0 - \langle r\rangle_0^2} = \sqrt{30a_0^2 - (5a_0)^2} = \sqrt{5}a_0. \tag{6.296}$$

Problem 6.8

The operators associated with the radial component of the momentum p_r and the radial coordinate r are denoted by $\hat{P}_r$ and $\hat{R}$, respectively. Their actions on a radial wave function $\psi_{(r)}$ are given by $\hat{P}_r\psi(\vec{r}) = -i\hbar(1/r)(\partial/\partial r)(r\psi(\vec{r}))$ and $\hat{R}\psi(\vec{r}) = r\psi(\vec{r})$.

(a) Find the commutator $[\hat{P}_r, \hat{R}]$ and $\Delta P_r \Delta r$, where $\Delta r = \sqrt{\langle \hat{R}^2 \rangle - \langle \hat{R} \rangle^2}$ and $\Delta P_r = \sqrt{\langle \hat{P}_r^2 \rangle - \langle \hat{P}_r \rangle^2}$.

(b) Show that $\hat{P}_r^2 = -(\hbar^2/r)(\partial^2/\partial r^2)r$.

Solution

(a) Since $\hat{R}\psi(\vec{r}) = r\psi(\vec{r})$ and

$$\hat{P}_r\psi(\vec{r}) = -i\hbar\frac{1}{r}\frac{\partial}{\partial r}(r\psi(\vec{r})) = -i\hbar\frac{1}{r}\psi(\vec{r}) - i\hbar\frac{\partial\psi(\vec{r})}{\partial r}, \tag{6.297}$$

and since

$$\hat{P}_r(\hat{R}\psi(\vec{r})) = -i\hbar\frac{1}{r}\frac{\partial}{\partial r}\left(r^2\psi(\vec{r})\right) = -2i\hbar\psi(\vec{r}) - i\hbar r\frac{\partial\psi(\vec{r})}{\partial r}, \tag{6.298}$$

the action of the commutator $[\hat{P}_r, \hat{R}]$ on a function $\psi(\vec{r})$ is given by

$$\begin{aligned}
[\hat{P}_r,\ \hat{R}]\psi(\vec{r}) &= -i\hbar\left[\frac{1}{r}\frac{\partial}{\partial r}r,\ \hat{R}\right]\psi(\vec{r}) = -i\hbar\frac{1}{r}\frac{\partial}{\partial r}\left(r^2\psi(\vec{r})\right) + i\hbar\frac{\partial}{\partial r}(r\psi(\vec{r})) \\
&= -2i\hbar\psi(\vec{r}) - i\hbar r\frac{\partial\psi(\vec{r})}{\partial r} + i\hbar\psi(\vec{r}) + i\hbar r\frac{\partial\psi(\vec{r})}{\partial r} \\
&= -i\hbar\psi(\vec{r}).
\end{aligned} \tag{6.299}$$

Thus, we have

$$[\hat{P}_r,\ \hat{R}] = -i\hbar. \tag{6.300}$$

Using the uncertainty relation for a pair of operators $\hat{A}$ and $\hat{B}$, $\Delta A \Delta B \geq \frac{1}{2}|\langle[\hat{A}, \hat{B}]\rangle|$, we can write

$$\Delta P_r \Delta r \geq \frac{1}{2}\left|\left\langle[\hat{P}_r, \hat{R}]\right\rangle\right|, \tag{6.301}$$

or

$$\Delta P_r \Delta r \geq \frac{\hbar}{2}. \tag{6.302}$$

(b) The action of $\hat{P}_r^2$ on $\psi(\vec{r})$ gives

$$\hat{P}_r^2\psi(\vec{r}) = -\hbar^2\frac{1}{r}\frac{\partial}{\partial r}\left[r\frac{1}{r}\frac{\partial}{\partial r}(r\psi)\right] = -\hbar^2\frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi(\vec{r})); \tag{6.303}$$

hence

$$\hat{P}_r^2 = -\hbar^2\frac{1}{r}\frac{\partial^2}{\partial r^2}(r). \tag{6.304}$$

Problem 6.9

Find the number of s bound states for a particle of mass m moving in a delta potential $V(r) = -V_0\delta(r - a)$ where $V_0 > 0$. Discuss the existence of bound states in terms of the size of a. Find the normalized wave function of the bound state(s).

Solution

The $l = 0$ radial equation can be obtained from (6.57):

$$\frac{d^2U_{n0}(r)}{dr^2} + \left[\frac{2mV_0}{\hbar^2}\delta(r-a) - k^2\right]U_{n0}(r) = 0, \tag{6.305}$$

where $U_{nl}(r) = U_{n0}(r) = rR_{n0}(r)$ and $k^2 = -2mE/\hbar^2$, since we are looking here at the bound states only, $E < 0$. The solutions of this equation are

$$U_{n0}(r) = \begin{cases} U_{n0_1}(r) = Ae^{kr} + Be^{-kr}, & 0 < r < a, \\ U_{n0_2}(r) = Ce^{-kr}, & r > a. \end{cases} \tag{6.306}$$

The energy eigenvalues can be obtained from the boundary conditions. As the wave function vanishes at $r = 0$, $U_{n0}(0) = 0$, we have $A + B = 0$ or $B = -A$; hence $U_{n0_1}(r) = D\sinh kr$:

$$U_{n0}(r) = D\sinh kr, \qquad 0 < r < a, \tag{6.307}$$

with $D = 2A$. The continuity condition at $r = a$ of $U_{n0}(r)$, $U_{n0_1}(a) = U_{n0_2}(a)$, leads to

$$D\sinh ka = Ce^{-ka}. \tag{6.308}$$

To obtain the discontinuity condition for the first derivative of $U_{n0}(r)$ at $r = a$, we need to integrate (6.305):

$$\lim_{\varepsilon\to a}\left[U'_{n0_2}(a+\varepsilon) - U'_{n0_1}(a-\varepsilon)\right] + \frac{2mV_0}{\hbar^2}U_{n0_2}(a) = 0 \tag{6.309}$$

or

$$-kCe^{-ka} - kD\cosh ka + \frac{2mV_0}{\hbar^2}Ce^{-ka} = 0. \tag{6.310}$$

Taking $Ce^{-ka} = D\sinh ka$, as given by (6.308), and substituting it into (6.310), we get

$$-k\sinh ka - k\cosh ka + \frac{2mV_0}{\hbar^2}\sinh ka = 0; \tag{6.311}$$

hence

$$\gamma\coth\gamma = \frac{2mV_0}{\hbar^2}a - \gamma, \tag{6.312}$$

where $\gamma = ka$.

The energy eigenvalues are given by the intersection of the curves $f(\gamma) = \gamma\coth\gamma$ and $g(\gamma) = 2mV_0a/\hbar^2 - \gamma$. As shown in Figure 6.7, if $a < \hbar^2/(2mV_0)$ then no bound state solution can exist, since the curves of $f(\gamma)$ and $g(\gamma)$ do not intersect. But if $a > \hbar^2/(2mV_0)$ the curves intersect only once; hence there is one bound state. We can summarize these results as follows:

$$a < \frac{\hbar^2}{2mV_0} \Longrightarrow \text{no bound states}, \tag{6.313}$$

$$a > \frac{\hbar^2}{2mV_0} \Longrightarrow \text{only one bound state}. \tag{6.314}$$

The radial wave function is given by

$$R_{n0}(r) = \frac{1}{r}U_{n0}(r) = \begin{cases} (D/r)\sinh kr, & 0 < r < a, \\ (C/r)e^{-kr}, & r > a. \end{cases} \tag{6.315}$$

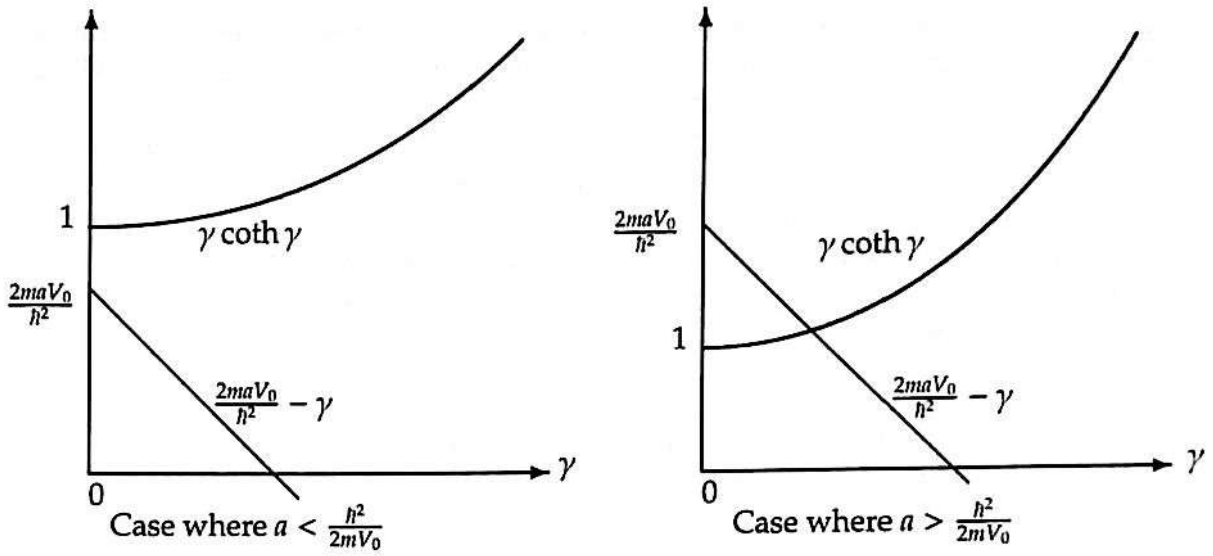


Figure 6.7: Graphical solutions of $f(\gamma) = g(\gamma)$, with $\gamma = ka$, $f(\gamma) = \gamma \coth \gamma$, and $g(\gamma) = 2mV_0a/\hbar^2 - \gamma$. If $a < \hbar^2/(2mV_0)$ there is no bound state. If $a > \hbar^2/(2mV_0)$ there is one bound state.

The normalization of this function yields

$$\begin{aligned} 1 &= \int_0^\infty r^2 R_{n0}^2(r)\,dr = \int_0^\infty U_{n0}^2(r)\,dr \\ &= D^2 \int_0^a \sinh^2 kr\,dr + C^2 \int_a^\infty e^{-2kr} dr = \frac{D^2}{2} \int_0^a [\cosh 2kr - 1]\,dr + \frac{C^2}{2k} e^{-2ka} \\ &= D^2 \left[\frac{1}{4k} \sinh 2ka - \frac{a}{2}\right] + \frac{C^2}{2k} e^{-2ka}. \end{aligned} \tag{6.316}$$

From (6.308) we have $Ce^{-ka} = D \sinh ka$, so we can rewrite this relation as

$$1 = D^2 \left[\frac{1}{4k} \sinh 2ka - \frac{a}{2}\right] + \frac{D^2}{2k} \sinh^2 ka = D^2 \left[\frac{\sinh 2ka + 2\sinh^2 ka}{4k} - \frac{a}{2}\right]; \tag{6.317}$$

hence

$$D = \frac{2\sqrt{k}}{\sqrt{\sinh 2ka + 2\sinh^2 ka - 2ak}}. \tag{6.318}$$

The normalized wave function is thus given by $\psi_{nlm}(r) = \psi_{n00}(r) = (1/\sqrt{4\pi}) R_{n0}(r)$ or

$$\psi_{n00}(r) = \frac{\sqrt{k}}{\sqrt{\pi \sinh 2ka + 2\pi \sinh^2 ka - 2\pi ak}} \begin{cases} (1/r)\sinh(kr), & 0 < r < a, \\ (1/r)\sinh(ka)e^{-k(r-a)}, & r > a. \end{cases} \tag{6.319}$$

Problem 6.10
Consider the $l = 0$ states of a bound system of two quarks having the same mass m and interacting via $V(r) = kr$.

(a) Using the Bohr model, find the speed, the radius, and the energy of the system in the case of circular orbits. Determine also the angular frequency of the radiation generated by a transition of the system from an energy state n to m.

(b) Solve the Schrödinger equation for the central potential $V(r) = kr$ for the two-quark system and find the expressions for the energy and the radial function $R_{nl}(r)$. Compare the energy with the value obtained in (a).

(c) Use the expressions derived in (a) and (b) to calculate the four lowest energy levels of a bottom–antibottom (bottomonium) quark system with $k = 15\,\text{GeV fm}^{-1}$; the mass–energy of a bottom quark is $mc^2 = 4.4\,\text{GeV}$.

Solution

(a) Consider the two quarks to move circularly, much like the electron and proton in a hydrogen atom; we can write the force between them as

$$\mu\frac{v^2}{r} = \frac{dV(r)}{dr} = k, \tag{6.320}$$

where $\mu = m/2$ is the reduced mass. From the Bohr quantization condition of the orbital angular momentum, we have

$$L = \mu v r = n\hbar, \tag{6.321}$$

Multiplying (6.320) by (6.321), we end up with $\mu^2 v^3 = n\hbar k$ which yields the speed of the relative motion of the two-quark system:

$$v_n = \left(\frac{n\hbar k}{\mu^2}\right)^{1/3}. \tag{6.322}$$

The radius can be obtained from (6.321), $r_n = n\hbar/(\mu v_n)$; using (6.322) this leads to

$$r_n = \left(\frac{n^2\hbar^2}{\mu k}\right)^{1/3}. \tag{6.323}$$

We can obtain the total energy of the relative motion by adding the kinetic and potential energies:

$$E_n = \frac{1}{2}\mu v_n^2 + k r_n = \frac{3}{2}\left(\frac{n^2\hbar^2 k^2}{\mu}\right)^{1/3}. \tag{6.324}$$

In deriving this we have used the relations for v_n and r_n as given by (6.322) and (6.323), respectively. The angular frequency of the radiation generated by a transition from n to m is given by

$$\omega_{nm} = \frac{E_n - E_m}{\hbar} = \frac{3}{2\hbar}\left(\frac{k^2}{\mu\hbar}\right)^{1/3}\left(n^{2/3} - m^{2/3}\right). \tag{6.325}$$

(b) The radial equation is given by (6.57):

$$-\frac{\hbar^2}{2\mu}\frac{d^2U_{nl}(r)}{dr^2} + \left[kr + \frac{l(l+1)\hbar^2}{2Mr^2}\right]U_{nl}(r) = E_n U_{nl}(r), \tag{6.326}$$

where $U_{nl}(r) = rR_{nl}(r)$. Since we are dealing with $l = 0$, we have

$$-\frac{\hbar^2}{2\mu}\frac{d^2U_{n0}(r)}{dr^2} + krU_{n0}(r) = E_nU_{n0}(r), \tag{6.327}$$

which can be reduced to

$$\frac{d^2U_{n0}(r)}{dr^2} - \frac{2\mu k}{\hbar^2}\left(r - \frac{E}{k}\right)U_{n0}(r) = 0. \tag{6.328}$$

Making the change of variable $x = (2\mu k/\hbar^2)^{1/3}(r - E/k)$, we can rewrite (6.328) as

$$\frac{d^2\phi_n(x)}{dx^2} - x\phi_n(x) = 0. \tag{6.329}$$

We have already studied the solutions of this equation in Chapter 4; they are given by the Airy functions $\mathrm{Ai}(x)$: $\phi(x) = B\mathrm{Ai}(x)$. The bound state energies result from the zeros of $\mathrm{Ai}(x)$. The boundary conditions on U_{nl} of (6.329) are $U_{nl}(r = 0) = 0$ and $U_{nl}(r \to +\infty) = 0$. The second condition is satisfied by the Airy functions, since $\mathrm{Ai}(x \to +\infty) = 0$. The first condition corresponds to $\phi[-(2\mu k/\hbar^2)^{1/3}E/k] = 0$ or to $\mathrm{Ai}[-(2\mu k/\hbar^2)^{1/3}E/k] = \mathrm{Ai}(R_n) = 0$, where R_n are the zeros of the Airy function.

The boundary condition $U_{nl}(r = 0) = 0$ then yields a discrete set of energy levels which can be expressed in terms of the Airy roots as follows:

$$\mathrm{Ai}\left[-\left(\frac{2\mu k}{\hbar^2}\right)^{1/3}\frac{E}{k}\right] = 0 \quad \Longrightarrow \quad -\left(\frac{2\mu k}{\hbar^2}\right)^{1/3}\frac{E_n}{k} = R_n; \tag{6.330}$$

hence

$$E_n = -\left(\frac{\hbar^2k^2}{2\mu}\right)^{1/3}R_n. \tag{6.331}$$

The radial function of the system is given by $R_{n0}(r) = (1/r)U_{n0}(r) = (B_n/r)\mathrm{Ai}(x)$ or

$$R_{n0}(r) = \frac{B_n}{r}\mathrm{Ai}(x) = \frac{B_n}{r}\mathrm{Ai}\left[\left(\frac{2\mu k}{\hbar^2}\right)^{1/3}r + R_n\right]. \tag{6.332}$$

The energy expression (6.331) has the same structure as the energy (6.324) derived from the Bohr model $E_n^B = \frac{3}{2}(n^2\hbar^2k^2/\mu)^{1/3}$; the ratio of the two expressions is

$$\frac{E_n}{E_n^B} = -\frac{2}{3}\frac{R_n}{(2n^2)^{1/3}}. \tag{6.333}$$

(c) In the following calculations we will be using $k = 15\,\mathrm{GeV\,fm^{-1}}$, $\mu c^2 = mc^2/2 = 2.2\,\mathrm{GeV}$, and $\hbar c = 197.3\,\mathrm{MeV\ fm}$. The values of the four lowest energy levels corresponding to the expression $E_n^B = \frac{3}{2}(n^2\hbar^2k^2/\mu)^{1/3}$, derived from the Bohr model, are

$$E_1^B = \frac{3}{2}\left(\frac{\hbar^2k^2}{\mu}\right)^{1/3} = 2.38\,\mathrm{GeV}, \qquad E_2^B = 2^{2/3}E_1^B = 3.77\,\mathrm{GeV}, \tag{6.334}$$

$$E_3^B = 3^{2/3}E_1^B = 4.95\,\mathrm{GeV}, \qquad E_4^B = 4^{2/3}E_1^B = 5.99\,\mathrm{GeV}. \tag{6.335}$$

Let us now calculate the exact energy levels. As mentioned in Chapter 4, the first few roots of the Airy function are given by $R_1 = -2.338$, $R_2 = -4.088$, $R_3 = -5.521$, $R_4 = -6.787$, so we can immediately obtain the first few energy levels:

$$E_1 = \left(\frac{\hbar^2 k^2}{2\mu}\right)^{1/3} R_1 = 2.94\,\text{GeV}, \qquad E_2 = \left(\frac{\hbar^2 k^2}{2\mu}\right)^{1/3} R_2 = 5.14\,\text{GeV}, \tag{6.336}$$

$$E_3 = \left(\frac{\hbar^2 k^2}{2\mu}\right)^{1/3} R_3 = 6.95\,\text{GeV}, \qquad E_4 = \left(\frac{\hbar^2 k^2}{2\mu}\right)^{1/3} R_4 = 8.54\,\text{GeV}. \tag{6.337}$$

Problem 6.11

Consider a system of two spinless particles of reduced mass μ that is subject to a finite, central potential well

$$V(r) = \begin{cases} -V_0, & 0 \le r \le a, \\ 0, & r > a, \end{cases}$$

where V_0 is positive. The purpose of this problem is to show how to find the minimum value of V_0 so that the potential well has one $l = 0$ bound state.

(a) Find the solution of the radial Schrödinger equation in both regions, $0 \le r \le a$ and $r > a$, in the case where the particle has zero angular momentum and its energy is located in the range $-V_0 < E < 0$.

(b) Show that the continuity condition of the radial function at $r = a$ can be reduced to a transcendental equation in E.

(c) Use this continuity condition to find the minimum values of V_0 so that the system has one, two, and three bound states.

(d) Obtain the results of (c) from a graphical solution of the transcendental equation derived in (b).

(e) Use the expression obtained in (c) to estimate a numerical value of V_0 for a deuteron nucleus with $a = 2 \times 10^{-15}$ m; a deuteron nucleus consists of a neutron and a proton.

Solution

(a) When $l = 0$ and $-V_0 < E < 0$ the radial equation (6.56),

$$-\frac{\hbar^2}{2\mu}\frac{d^2U_{nl}(r)}{dr^2} + \left[\frac{l(l+1)\hbar^2}{2\mu r^2} + V(r)\right]U_{nl}(r) = E_n U_{nl}(r), \tag{6.338}$$

can be written inside the well, call it region (1), as

$$U_n''(r)_1 + k_1^2 U_n(r)_1 = 0, \qquad 0 \le r \le a, \tag{6.339}$$

and outside the well, call it region (2), as

$$U_n''(r)_2 - k_2^2 U_n(r)_2 = 0, \qquad r > a, \tag{6.340}$$

where $U_n''(r) = d^2U_n(r)/dr^2$, $U_n(r)_1 = rR_n(r)_1$, $U_n(r)_2 = rR_n(r)_2$, $k_1 = \sqrt{2\mu(V_0 + E)/\hbar^2}$ and $k_2 = \sqrt{-2\mu E/\hbar^2}$. Since $U_n(r)_1$ must vanish at $r = 0$, while $U_n(r)_2$ has to be finite at $r \longrightarrow \infty$, the respective solutions of (6.339) and (6.340) are given by

$$U_n(r)_1 = A\sin(k_1 r), \qquad 0 \le r \le a, \tag{6.341}$$

$$U_n(r)_2 = Be^{-k_2 r}, \qquad r > a. \tag{6.342}$$

The corresponding radial functions are

$$R_n(r)_1 = A\frac{\sin(k_1 r)}{r}, \qquad R_n(r)_2 = B\frac{e^{-k_2 r}}{r}. \tag{6.343}$$

(b) Since the logarithmic derivative of the radial function is continuous at $r = a$, we can write

$$\frac{R_n'(a)_1}{R_n(a)_1} = \frac{R_n'(a)_2}{R_n(a)_2}. \tag{6.344}$$

From (6.343) we have

$$\frac{R_n'(a)_1}{R_n(a)_1} = k_1 \cot(k_1 a) - \frac{1}{a}, \qquad \frac{R_n'(a)_2}{R_n(a)_2} = -k_2 - \frac{1}{a}. \tag{6.345}$$

Substituting (6.345) into (6.344) we obtain

$$-k_1 \cot(k_1 a) = k_2 \tag{6.346}$$

or

$$\sqrt{\frac{2\mu}{\hbar^2}(V_0 + E)}\cot\left[\sqrt{\frac{2\mu}{\hbar^2}(V_0 + E)}a\right] = -\sqrt{-\frac{2\mu E}{\hbar^2}}, \tag{6.347}$$

since $k_1 = \sqrt{2\mu(V_0 + E)/\hbar^2}$ and $k_2 = \sqrt{-2\mu E/\hbar^2}$.

(c) In the limit $E \to 0$, the system has very few bound states; in this limit, equation (6.347) becomes

$$\sqrt{\frac{2\mu V_0}{\hbar^2}}\cot\left(\sqrt{\frac{2\mu V_0}{\hbar^2}}a\right) = 0, \tag{6.348}$$

which leads to $a\sqrt{2\mu V_{0n}/\hbar^2} = (2n + 1)\pi/2$; hence

$$V_{0n} = \frac{\pi^2\hbar^2}{8\mu a^2}(2n + 1)^2, \qquad n = 0, 1, 2, 3, \ldots. \tag{6.349}$$

Thus, the minimum values of V_0 corresponding to one, two, and three bound states are respectively

$$V_{00} = \frac{\pi^2\hbar^2}{8\mu a^2}, \qquad V_{01} = \frac{9\pi^2\hbar^2}{8\mu a^2}, \qquad V_{02} = \frac{25\pi^2\hbar^2}{8\mu a^2}. \tag{6.350}$$

(d) Using the notation $\alpha = ak_1$ and $\beta = ak_2$ we can, on the one hand, write

$$\alpha^2 + \beta^2 = \frac{2\mu a^2 V_0}{\hbar^2}, \tag{6.351}$$

and, on the other hand, reduce the transcendental equation (6.346) to

$$-\alpha \cot\alpha = \beta, \tag{6.352}$$

since $k_1 = \sqrt{2\mu(V_0 + E)/\hbar^2}$ and $k_2 = \sqrt{-2\mu E/\hbar^2}$.

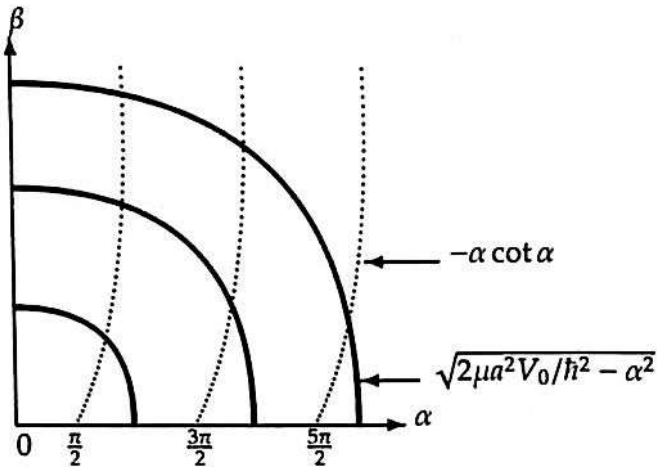


Figure 6.8: Graphical solutions for the finite, spherical square well potential: they are given by the intersection of the circle $\alpha^2+\beta^2=2\mu a^2V_0/\hbar^2$ with the curve of $-\alpha\cot\alpha$, where $\alpha^2=2\mu a^2(V_0+E)/\hbar^2$ and $\beta^2=-2\mu a^2E/\hbar^2$, with $-V_0<E<0$.

As shown in Figure 6.8, when $\pi/2<\alpha<3\pi/2$, which in the limit of $E\to 0$ leads to

$$\frac{\pi^2\hbar^2}{8\mu a^2}<V_0<\frac{9\pi^2\hbar^2}{8\mu a^2}, \tag{6.353}$$

there exists only one bound state, since the circle intersects only once with the curve $-\alpha\cot\alpha$. Similarly, there are two bound states if $3\pi/2<\alpha<5\pi/2$ or

$$\frac{9\pi^2\hbar^2}{8\mu a^2}<V_0<\frac{25\pi^2\hbar^2}{8\mu a^2}, \tag{6.354}$$

and three bound states if $5\pi/2<\alpha<7\pi/2$:

$$\frac{25\pi^2\hbar^2}{8\mu a^2}<V_0<\frac{49\pi^2\hbar^2}{8\mu a^2}. \tag{6.355}$$

(e) Since $m_pc^2\simeq 938\,\text{MeV}$ and $m_nc^2\simeq 940\,\text{MeV}$, the reduced mass of the deuteron is given by $\mu c^2=(m_pc^2)(m_nc^2)/(m_pc^2+m_nc^2)\simeq 469.5\,\text{MeV}$. Since $a=2\times10^{-15}\,\text{m}$ the minimum value of V_0 corresponding to one bound state is

$$V_0=\frac{\pi^2\hbar^2}{8\mu a^2}=\frac{\pi^2(\hbar c)^2}{8(\mu c^2)a^2}=\frac{\pi^2(197\,\text{MeV fm})^2}{8(469.5\,\text{MeV})(2\times10^{-15}\,\text{m})^2}\simeq 25.5\,\text{MeV}. \tag{6.356}$$

Problem 6.12
Calculate $\langle nl\,|\,\hat{P}^4\,|\,nl\rangle$ in a stationary state $|\,nl\rangle$ of the hydrogen atom.

Solution

To calculate $\langle nl \mid \hat{P}^4 \mid nl\rangle$ we may consider expressing $\hat{P}^4$ in terms of the hydrogen's Hamiltonian. Since $\hat{H} = \hat{P}^2/(2m_e) - e^2/r$ we have $\hat{P}^2 = 2m_e(\hat{H} + e^2/r)$; hence

$$\begin{aligned}\langle nl \mid \hat{P}^4 \mid nl\rangle &= (2m_e)^2 \left\langle nl \left| \left(\hat{H} + \frac{e^2}{r}\right)^2 \right| nl \right\rangle \\ &= (2m_e)^2 \left\langle nl \left| \hat{H}^2 + \hat{H}\frac{e^2}{r} + \frac{e^2}{r}\hat{H} + \frac{e^4}{r^2} \right| nl \right\rangle \\ &= (2m_e)^2 \left[E_n^2 + E_n \left\langle nl \left| \frac{e^2}{r} \right| nl \right\rangle + \left\langle nl \left| \frac{e^2}{r} \right| nl \right\rangle E_n + \left\langle nl \left| \frac{e^4}{r^2} \right| nl \right\rangle \right],\end{aligned} \tag{6.357}$$

where we have used the fact that $\mid nl\rangle$ is an eigenstate of $\hat{H}$: $\hat{H} \mid nl\rangle = E_n \mid nl\rangle$ with $E_n = -e^2/(2a_0n^2) = -13.6\,\text{eV}/n^2$. The expectation values of $1/r$ and $1/r^2$ are given by (6.182) and (6.183), $\langle nl|r^{-1}|nl\rangle = 1/(n^2a_0)$ and $\langle nl|r^{-2}|nl\rangle = 2/[n^3(2l+1)a_0^2]$; we can thus rewrite (6.357) as

$$\begin{aligned}\langle nl \mid \hat{P}^4 \mid nl\rangle &= (2m_e)^2 \left[E_n^2 + 2E_n \left\langle nl \left| \frac{e^2}{r} \right| nl \right\rangle + \left\langle nl \left| \frac{e^4}{r^2} \right| nl \right\rangle \right] \\ &= (2m_eE_n)^2 \left[1 + \frac{2e^2}{E_n}\frac{1}{n^2\,a_0} + \frac{e^4}{E_n^2}\frac{2}{n^3(2l\,+1)a_0^2} \right] \\ &= (2m_eE_n)^2 \left[1 - 4 + \frac{8n}{2l+1} \right];\end{aligned} \tag{6.358}$$

in deriving the last relation we have used $E_n = -e^2/(2a_0n^2)$. Now, since $a_0 = \hbar^2/(m_ee^2)$, the energy E_n becomes $E_n = -e^2/(2a_0n^2) = -m_ee^4/(2\hbar^2n^2)$ which, when inserted into (6.358), leads to

$$\boxed{\langle nl \mid \hat{P}^4 \mid nl\rangle = \frac{m_e^4e^8}{\hbar^4n^4}\left[\frac{8n}{2l+1} - 3\right].} \tag{6.359}$$

6.7 Exercises

Exercise 6.1

A spinless particle of mass m is confined to move in the xy plane under the influence of a harmonic oscillator potential $\hat{V}(x,y) = \frac{1}{2}m\omega^2\left(\hat{x}^2 + \hat{y}^2\right)$ for all values of x and y.

(a) Show that the Hamiltonian $\hat{H}$ of this particle can be written as a sum of two familiar one-dimensional Hamiltonians, $\hat{H}_x$ and $\hat{H}_y$. Then show that $\hat{H}$ commutes with $\hat{L}_z = \hat{X}\hat{P}_y - \hat{Y}\hat{P}_x$.

(b) Find the expression for the energy levels $E_{n_xn_y}$.

(c) Find the energies of the four lowest states and their corresponding degeneracies.

(d) Find the degeneracy g_n of the nth excited state as a function of the quantum number n ($n = n_x + n_y$).

(e) If the state vector of the nth excited state is $|n\rangle = |n_x\rangle|n_y\rangle$ or

$$\langle xy|n\rangle = \langle x|n_x\rangle\langle y|n_y\rangle = \psi_{n_x}(x)\psi_{n_y}(y),$$

calculate the expectation value of the operator $\hat{A} = \hat{x}^4 + \hat{y}^2$ in the state $|n\rangle$ as a function of the quantum numbers n_x and n_y.

Exercise 6.2
A particle of mass m moves in the xy plane in the potential

$$V(x,y) = \begin{cases} \frac{1}{2}m\omega^2 y^2 & \text{for all } y \text{ and } 0 < x < a, \\ +\infty, & \text{elsewhere.} \end{cases}$$

(a) Write down the time-independent Schrödinger equation for this particle and reduce it to a set of familiar one-dimensional equations.
(b) Find the normalized eigenfunctions and the eigenenergies.

Exercise 6.3
A particle of mass m moves in the xy plane in a two-dimensional rectangular well

$$V(x,y) = \begin{cases} 0, & 0 < x < a, \quad 0 < y < b, \\ +\infty, & \text{elsewhere.} \end{cases}$$

By reducing the time-independent Schrödinger equation to a set of more familiar one-dimensional equations, find the normalized wave functions and the energy levels of this particle.

Exercise 6.4
Consider an anisotropic three-dimensional harmonic oscillator potential

$$V(x,y,z) = \frac{1}{2}m(\omega_x^2 x^2 + \omega_y^2 y^2 + \omega_z^2 z^2).$$

(a) Evaluate the energy levels in terms of ω_x, ω_y, and ω_z.
(b) Calculate $[\hat{H}, \hat{L}_z]$. Do you expect the wave functions to be eigenfunctions of $\hat{L}^2$?
(c) Find the three lowest levels for the case $\omega_x = \omega_y = 2\omega_z/3$, and determine the degeneracy of each level.

Exercise 6.5
Consider a spinless particle of mass m which is confined to move under the influence of a three-dimensional potential

$$\hat{V}(x,y,z) = \begin{cases} 0 & \text{for } 0 < x < a,\ 0 < y < a,\ 0 < z < b, \\ +\infty & \text{elsewhere.} \end{cases}$$

(a) Find the expression for the energy levels $E_{n_x n_y n_z}$ and their corresponding wave functions.
(b) If $a = 2b$ find the energies of the five lowest states and their degeneracies.

Exercise 6.6
A particle of mass m moves in the three-dimensional potential

$$V(x,y,z) = \begin{cases} \frac{1}{2}m\omega^2 z^2 & \text{for } 0 < x < a,\ 0 < y < a, \text{ and } z > 0 \\ +\infty, & \text{elsewhere.} \end{cases}$$

(a) Write down the time-independent Schrödinger equation for this particle and reduce it to a set of familiar one-dimensional equations; then find the normalized wave function $\psi_{n_x n_y n_z}(x, y, z)$.

(b) Find the allowed eigenenergies of this particle and show that they can be written as: $E_{n_x n_y n_z} = E_{n_x n_y} + E_{n_z}$.

(c) Find the four lowest energy levels in the xy plane (i.e., $E_{n_x n_y}$) and their corresponding degeneracies.

Exercise 6.7

A particle of mass m moves in the potential $V(x, y, z) = V_1(x, y) + V_2(z)$ where

$$V_1(x, y) = \frac{1}{2}m\omega^2\left(x^2 + y^2\right), \qquad V_2(z) = \begin{cases} 0, & 0 \le z \le a, \\ +\infty, & \text{elsewhere.} \end{cases}$$

(a) Calculate the energy levels and the wave function of this particle.

(b) Let us now turn off $V_2(z)$ (i.e., m is subject only to $V_1(x, y)$). Calculate the degeneracy g_n of the nth energy level (note that $n = n_x + n_y$).

Exercise 6.8

Consider a muonic atom which consists of a nucleus that has Z protons (no neutrons) and a negative muon moving around it; the muon's charge is $-e$ and its mass is 207 times the mass of the electron, $m_{\mu^-} = 207m_e$. For a muonic atom with $Z = 6$, calculate

(a) the radius of the first Bohr orbit,

(b) the energy of the ground, first, and second excited states, and

(c) the frequency associated with the transitions $n_i = 2 \rightarrow n_f = 1$, $n_i = 3 \rightarrow n_f = 1$, and $n_i = 3 \rightarrow n_f = 2$.

Exercise 6.9

A hydrogen atom has the wave function $\Psi_{nlm}(\vec{r})$, where $n = 4$, $l = 3$, $m = 3$.

(a) What is the magnitude of the orbital angular momentum of the electron around the proton?

(b) What is the angle between the orbital angular momentum vector and the z-axis? Can this angle be reduced by changing n or m if l is held constant? What is the physical significance of this result?

(c) Sketch the shapes of the radial function and of the probability of finding the electron a distance r from the proton.

Exercise 6.10

An electron in a hydrogen atom is in the energy eigenstate

$$\psi_{2,1,-1}(r, \theta, \varphi) = Nre^{-r/2a_0}Y_{1,-1}(\theta, \varphi).$$

(a) Find the normalization constant, N.

(b) What is the probability per unit volume of finding the electron at $r = a_0$, $\theta = 45°$, $\varphi = 60°$?

(c) What is the probability per unit radial interval (dr) of finding the electron at $r = 2a_0$? (One must take an integral over θ and φ at $r = 2a_0$.)

(d) If the measurements of $\hat{L}^2$ and $\hat{L}_z$ were carried out, what will be the results?

Exercise 6.11
Consider a hydrogen atom which is in its ground state; the ground state wave function is given by

$$\Psi(r,\theta,\varphi) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0},$$

where a_0 is the Bohr radius.

(a) Find the most probable distance between the electron and the proton when the hydrogen atom is in its ground state.

(b) Find the average distance between the electron and the proton.

Exercise 6.12
Consider a hydrogen atom whose state at time $t = 0$ is given by

$$\Psi(\vec{r},0) = \frac{1}{\sqrt{2}}\phi_{300}(\vec{r}) + \frac{1}{\sqrt{3}}\phi_{311}(\vec{r}) + \frac{1}{\sqrt{6}}\phi_{322}(\vec{r}).$$

(a) What is the time-dependent wave function?

(b) If a measurement of the energy were carried out, what values could be found and with what probabilities?

(c) Repeat part (b) for $\hat{L}^2$ and $\hat{L}_z$. That is, if a measurement of $\hat{L}^2$ and $\hat{L}_z$ were carried out, what values could be found and with what probabilities?

Exercise 6.13
The wave function of an electron in a hydrogen atom is given by

$$\psi_{21m_l m_s}(r,\theta,\varphi) = R_{21}(r)\left[\frac{1}{\sqrt{3}}Y_{10}(\theta,\varphi)\left|\frac{1}{2},\frac{1}{2}\right\rangle + \sqrt{\frac{2}{3}}Y_{11}(\theta,\varphi)\left|\frac{1}{2},-\frac{1}{2}\right\rangle\right],$$

where $\left|\frac{1}{2},\pm\frac{1}{2}\right\rangle$ are the spin state vectors.

(a) Is this wave function an eigenfunction of $\hat{J}_z$, the z-component of the electron's total angular momentum? If yes, find the eigenvalue. (*Hint:* For this, you need to calculate $\hat{J}_z\psi_{21m_l m_s}$.)

(b) If you measure the z-component of the electron's spin angular momentum, what values will you obtain? What are the corresponding probabilities?

(c) If you measure $\hat{\vec{J}}^2$, what values will you obtain? What are the corresponding probabilities?

Exercise 6.14
Consider a hydrogen atom whose state at time $t = 0$ is given by

$$\Psi(\vec{r},0) = A\phi_{200}(\vec{r}) + \frac{1}{\sqrt{5}}\phi_{311}(\vec{r}) + \frac{1}{\sqrt{3}}\phi_{422}(\vec{r}),$$

where A is a normalization constant.

(a) Find A so that the state is normalized.

(b) Find the state of this atom at any later time t.

(c) If a measurement of the energy were carried out, what values would be found and with what probabilities?

(d) Find the mean energy of the atom.

Exercise 6.15
Calculate the width of the probability density distribution for r for the hydrogen atom in its ground state: $\Delta r = \sqrt{\langle r^2 \rangle_{10} - \langle r \rangle_{10}^2}$.

Exercise 6.16
Consider a hydrogen atom whose wave function is given at time $t = 0$ by

$$\psi(\vec{r}, 0) = \frac{A}{\sqrt{\pi}} \left(\frac{1}{a_0}\right)^{3/2} e^{-r/a_0} + \frac{1}{\sqrt{2\pi}} \left(\frac{z - \sqrt{2}x}{r}\right) R_{21}(r),$$

where A is a real constant, a_0 is the Bohr radius, and $R_{21}(r)$ is the radial wave function: $R_{21}(r) = 1/\sqrt{6}(1/a_0)^{3/2}(r/2a_0)e^{-r/2a_0}$.

(a) Write down $\psi(\vec{r}, 0)$ in terms of $\sum_{nlm} \phi_{nlm}(\vec{r})$ where $\phi_{nlm}(\vec{r})$ is the hydrogen wave function $\phi_{nlm}(\vec{r}) = R_{nl}(r)Y_{lm}(\theta, \varphi)$.

(b) Find A so that $\psi(\vec{r}, 0)$ is normalized. (Recall that $\int \phi^*_{n'l'm'}(\vec{r})\phi_{nlm}(\vec{r})d^3r = \delta_{n',n}\delta_{l',l}\delta_{m',m}$.)

(c) Write down the wave function $\psi(\vec{r}, t)$ at any later time t.

(d) Is $\psi(\vec{r}, 0)$ an eigenfunction of $\hat{\vec{L}}^2$ and $\hat{L}_z$? If yes, what are the eigenvalues?

(e) If a measurement of the energy is made, what value could be found and with what probability?

(f) What is the probability that a measurement of $\hat{L}_z$ yields $1\hbar$?

(g) Find the mean value of r in the state $\psi(\vec{r}, 0)$.

Exercise 6.17
Consider a pendulum undergoing small harmonic oscillations (with angular frequency $\omega = \sqrt{g/l}$, where g is the acceleration due to gravity and l is the length of the pendulum). Show that the quantum energy levels and the corresponding degeneracies of the pendulum are given by $E_n = (n + 1)\hbar\omega$ and $g_n = n + 1$, respectively.

Exercise 6.18
Consider a proton that is trapped inside an infinite central potential well

$$V(r) = \begin{cases} -V_0, & 0 < r < a, \\ +\infty, & r \geq a, \end{cases}$$

where $V_0 = 5104.34$ MeV and $a = 10$ fm.

(a) Find the energy and the (normalized) radial wave function of this particle for the s states (i.e., $l = 0$).

(b) Find the number of bound states that have energies lower than zero; you may use the values $mc^2 = 938$ MeV and $\hbar c = 197$ MeV fm.

(c) Calculate the energies of the levels that lie just below and just above the zero-energy level; express your answer in MeV.

Exercise 6.19
Consider the function $\psi(\vec{r}) = -A(x + iy)e^{-r/2a_0}$, where a_0 is the Bohr radius and A is a real constant.

(a) Is $\psi(\vec{r})$ an eigenfunction to $\hat{\vec{L}}^2$ and $\hat{L}_z$? If yes, write $\psi(\vec{r})$ in terms of $R_{nl}(r)Y_{lm}(\theta, \varphi)$ and find the values of the quantum numbers n, m, l; $R_{nl}(r)$ are the radial wave functions of the hydrogen atom.

(b) Find the constant A so that $\psi(\vec{r})$ is normalized.

(c) Find the mean value of r and the most probable value of r in this state.

Exercise 6.20

The wave function of a hydrogen-like atom at time $t = 0$ is

$$\Psi(\vec{r}, 0) = \frac{1}{\sqrt{11}}\left[\sqrt{3}\psi_{2,1,-1}(\vec{r}) - \psi_{2,1,0}(\vec{r}) + \sqrt{5}\psi_{2,1,1}(\vec{r}) + \sqrt{2}\psi_{3,1,1}(\vec{r})\right],$$

where $\psi_{nlm}(\vec{r})$ is a normalized eigenfunction (i.e., $\psi_{nlm}(\vec{r}) = R_{nl}(r)Y_{lm}(\theta, \varphi)$).

(a) What is the time-dependent wave function?

(b) If a measurement of energy is made, what values could be found and with what probabilities?

(c) What is the probability for a measurement of $\hat{L}_z$ which yields $-1\hbar$?

Exercise 6.21

Using the fact that the radial momentum operator is given by $\hat{p}_r = -i\hbar\frac{1}{r}\frac{\partial}{\partial r}r$, calculate the commutator $[\hat{r}, \hat{p}_r]$ between the position operator, $\hat{r}$, and the radial momentum operator.

Exercise 6.22

Calculate $\Delta r \Delta p_r$ with respect to the state

$$\psi_{2,1,0}(\vec{r}) = \frac{1}{\sqrt{6}}\left(\frac{1}{a_0}\right)^{3/2}\frac{r}{2a_0}e^{-r/2a_0}\,Y_{1,0}(\theta, \varphi),$$

and verify that $\Delta r \Delta p_r$ satisfies the Heisenberg uncertainty principle.

Chapter 7

Rotations and Addition of Angular Momenta

In this chapter we deal with rotations, the properties of addition of angular momenta, and the properties of tensor operators.

7.1 Rotations in Classical Physics

A rotation is defined by an angle of rotation and an axis about which the rotation is performed. Knowing the rotation matrix R, we can determine how vectors transform under rotations; in a three-dimensional space, a vector $\vec{A}$ becomes $\vec{A}'$ when rotated: $\vec{A}' = R\vec{A}$. Rotations can be studied from two equivalent perspectives: *active* and *passive* rotations. In the following two subsections, we are going to offer a brief treatment of both of them.

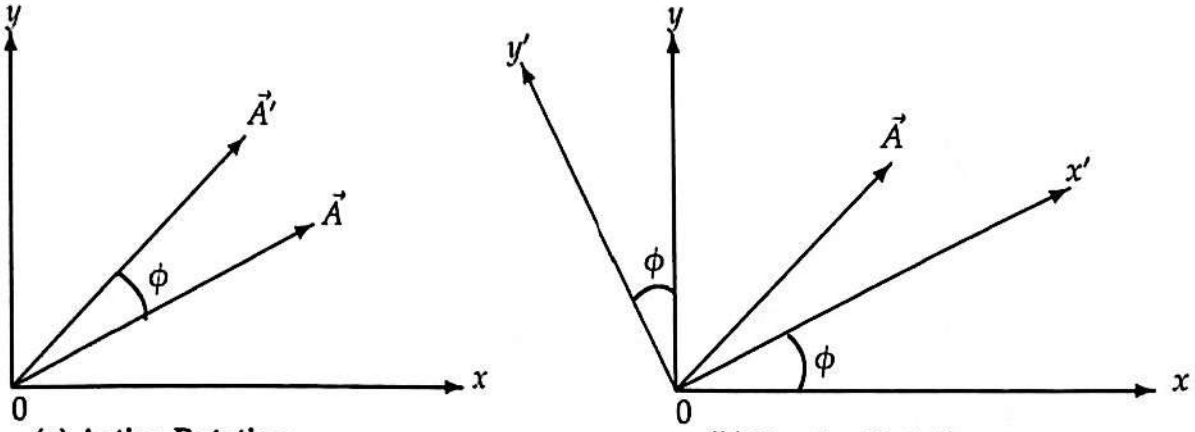
7.1.1 Active Rotations

For a simple illustration, consider a vector $\vec{A}$ in the $xy-$plane and examine its rotation about the $z-$axis. *In an active rotation, the coordinate system remains fixed while we rotate the vector.* For instance, when we rotate counterclockwise a vector $\vec{A} = A_x\hat{\imath} + A_y\hat{\jmath}$ through an angle ϕ about the $z-$axis it becomes $\vec{A}' = A'_x\hat{\imath} + A'_y\hat{\jmath}$; the magnitude of $\vec{A}$ remains unchanged under this rotation. As illustrated in Fig. 7.1-(a), we can show that this rotation transforms the coordinates of $\vec{A}$ as follows

$$\begin{pmatrix} A'_x \\ A'_y \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} A_x \\ A_y \end{pmatrix}. \tag{7.1}$$

We can easily generalize this case to a counterclockwise rotation of a more general, three-dimensional vector $\vec{A} = A_x\hat{\imath} + A_y\hat{\jmath} + A_z\hat{k}$ through an angle ϕ about the $z-$axis. Clearly, the $z-$component A_z will not be affected by this rotation; hence, it is easy to see that the vector $\vec{A}$

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.



(a) Active Rotation **(b) Passive Rotation**

Figure 7.1: **(a) Active rotation:** Keeping the xy–coordinate system fixed, we rotate the vector $\vec{A} = A_x\hat{\imath} + A_y\hat{\jmath}$ counterclockwise through an angle ϕ about the z–axis into the new vector $\vec{A}' = A'_x\hat{\imath} + A'_y\hat{\jmath}$; **(b)Passive Rotation:** Keeping the vector $\vec{A}$ intact, we rotate the xy–system counterclockwise through an angle ϕ about the z–axis into the new $x'y'$–system; vector $\vec{A}$ is expressed by $\vec{A} = A_x\hat{\imath} + A_y\hat{\jmath}$ and $\vec{A} = A'_x\hat{\imath}' + A'_y\hat{\jmath}'$ in the xy and $x'y'$ systems, respectively.

transforms like

$$\begin{pmatrix} A'_x \\ A'_y \\ A'_z \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix}. \tag{7.2}$$

We can write this expression as

$$\vec{A}' = R_z(\phi)\vec{A}, \tag{7.3}$$

where the rotation matrix is given by

$$R_z(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.4}$$

Similarly, we can ascertain that the rotation matrices corresponding to active counterclockwise rotations about the x– and y– axes are given by

$$R_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}, \qquad R_y(\phi) = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix}. \tag{7.5}$$

7.1.2 Passive Rotations

In a passive rotation, the vector remains fixed while we rotate the coordinate system. That is, and as shown in Fig. 7.1-(b), keeping the vector $\vec{A}$ fixed, when we rotate the xy–coordinate system counterclockwise through an angle ϕ about the z–axis, we end up with the $x'y'$–system. Vector $\vec{A}$ is expressed in the xy–system by $\vec{A} = A_x\hat{\imath} + A_y\hat{\jmath}$ and in the $x'y'$–system by $\vec{A} = A'_x\hat{\imath}' + A'_y\hat{\jmath}'$.

We can easily show that the coordinates of $\vec{A}$ in the $xy-$ and $x'y'-$coordinate systems, as a result of this passive rotation about the $z-$axis, are related by

$$\begin{pmatrix} A'_x \\ A'_y \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} A_x \\ A_y \end{pmatrix}. \tag{7.6}$$

Again, we can generalize this case to three-dimensions and easily show that the coordinates of $\vec{A}$ in the xyz and $x'y'z-$coordinate systems, as a result of this passive, counterclockwise rotation about the $z-$axis, are related by

$$\begin{pmatrix} A'_x \\ A'_y \\ A'_z \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \tag{7.7}$$

or $\vec{A}' = R_z(\phi)\vec{A}$, where

$$R_z(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.8}$$

Similarly, the rotation matrices corresponding to passive, counterclockwise rotations about the $x-$ and $y-$ axes are given by

$$R_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{pmatrix}, \qquad R_y(\phi) = \begin{pmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{pmatrix}. \tag{7.9}$$

Remarks

- Active and passive rotations can be connected to the two representations or pictures of quantum mechanics: the Heisenberg and Schrödinger pictures, which we are going to study in Chapter 10. *Active rotations* are analogous to the *Heisenberg picture* where state vectors are completely frozen in time (or unchanged by transformations), whereas operators evolve in time (or are changed by transformations). However, *passive rotations* are analogous to the *Schrödinger picture* in which state vectors depend on (or evolve in) time, whereas operators do not. That is, in active rotations, we rotate (or transform) the physical system while we hold the coordinate system fixed; in passive rotations, however, we keep the physical system fixed and rotate the coordinate system.

- Since they do not depend on the orientation of the coordinate system in which they are described, the laws of physics are invariant under rotations. For instance, the distance from any point to the origin (of the coordinate system) does not change under rotations of the coordinate system; that is, the length of a vector is invariant under coordinate rotations. To see this, using Eq (7.7), we can easily show that the square of the rotated vector $\vec{A}'$ is equal to that of $\vec{A}$:

$$\begin{aligned} A'^2_x &= (A_x\cos\phi + A_y\sin\phi)^2 = A_x^2\cos^2\phi + 2A_xA_y\cos\phi\sin\phi + A_y^2\sin^2\phi, \\ A'^2_y &= (-A_x\sin\phi + A_y\cos\phi)^2 = A_x^2\sin^2\phi - 2A_xA_y\sin\phi\cos\phi + A_y^2\cos^2\phi, \\ A'^2_z &= A_z^2, \end{aligned}$$

which lead at once to $\vec{A}'^{\,2} = \vec{A}^{\,2}$:

$$\begin{aligned} A'^2_x + A'^2_y + A'^2_z &= A_x^2(\cos^2\phi + \sin^2\phi) + 2A_xA_y(\cos\phi\sin\phi - \sin\phi\cos\phi) \\ &\quad + A_y^2(\sin^2\phi + \cos^2\phi) + A_z^2 \\ &= A_x^2 + A_y^2 + A_z^2. \end{aligned} \tag{7.10}$$

Using the same approach, we can also prove that the scalar product between any two vectors is invariant under rotations of the coordinate system; this is expected anyway since the scalar product is a scalar. Moreover, using Eq (7.7), we can easily verify the invariance of the scalar product, $\vec{A}' \cdot \vec{B}' = \vec{A} \cdot \vec{B}$:

$$\begin{aligned} A'_xB'_x &= (A_x\cos\phi + A_y\sin\phi)(B_x\cos\phi + B_y\sin\phi) \\ &= A_xB_x\cos^2\phi + A_xB_y\cos\phi\sin\phi + A_yB_x\sin\phi\cos\phi + A_yB_y\sin^2\phi, \\ A'_yB'_y &= (-A_x\sin\phi + A_y\cos\phi)(-B_x\sin\phi + B_y\cos\phi) \\ &= A_xB_x\sin^2\phi - A_xB_y\sin\phi\cos\phi - A_yB_x\cos\phi\sin\phi + A_yB_y\cos^2\phi, \\ A'_zB'_z &= A_zB_z, \end{aligned}$$

which lead to:

$$\begin{aligned} A'_xB'_x + A'_yB'_y + A'_zB'_z &= A_xB_x(\cos^2\phi + \sin^2\phi) + A_xB_y(\cos\phi\sin\phi - \sin\phi\cos\phi) \\ &\quad + A_yB_x(\sin\phi\cos\phi - \cos\phi\sin\phi) + A_y^2(\sin^2\phi + \cos^2\phi) + A_zB_z \\ &= A_xB_x + A_yB_y + A_zB_z. \end{aligned} \tag{7.11}$$

7.1.3 Properties of Rotation Matrices

We are going to illustrate the properties of rotation matrices on *active* rotations; the matrices of passive rotations obey the same properties as those of active rotations. From classical physics we know that, while rotations about the same axis commute, rotations about different axes do not. From (7.5) we can verify that $R_x(\phi)R_y(\phi) \neq R_y(\phi)R_x(\phi)$. In fact, using (7.5) we have

$$R_x(\phi)R_y(\phi) = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ \sin^2\phi & \cos\phi & -\cos\phi\sin\phi \\ -\cos\phi\sin\phi & \sin\phi & \cos^2\phi \end{pmatrix}, \tag{7.12}$$

$$R_y(\phi)R_x(\phi) = \begin{pmatrix} \cos\phi & \sin^2\phi & \cos\phi\sin\phi \\ 0 & \cos\phi & -\sin\phi \\ -\sin\phi & \cos\phi\sin\phi & \cos^2\phi \end{pmatrix}; \tag{7.13}$$

hence, $R_x(\phi)R_y(\phi) - R_y(\phi)R_x(\phi)$ is given by

$$\begin{pmatrix} 0 & -\sin^2\phi & \sin\phi - \cos\phi\sin\phi \\ \sin^2\phi & 0 & \sin\phi - \cos\phi\sin\phi \\ \sin\phi - \cos\phi\sin\phi & \sin\phi - \cos\phi\sin\phi & 0 \end{pmatrix}. \tag{7.14}$$

In the case of infinitesimal rotations of angle δ about the $x-$, $y-$, $z-$ axes, and using the approximations $\cos\delta \simeq 1 - \delta^2/2$ and $\sin\delta \simeq \delta$, we can reduce (7.14) to

$$R_x(\delta)R_y(\delta) - R_y(\delta)R_x(\delta) = \begin{pmatrix} 0 & -\delta^2 & 0 \\ \delta^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{7.15}$$

Now, from (7.4), we have

$$R_z(\delta) = \begin{pmatrix} 1-\frac{\delta^2}{2} & -\delta & 0 \\ \delta & 1-\frac{\delta^2}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \Longrightarrow R_z(\delta^2) = \begin{pmatrix} 1-\frac{\delta^4}{2} & -\delta^2 & 0 \\ \delta^2 & 1-\frac{\delta^4}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \simeq \begin{pmatrix} 1 & -\delta^2 & 0 \\ \delta^2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.16}$$

Using this expression of $R_z(\delta^2)$, we can rewrite (7.15) as

$$R_x(\delta)R_y(\delta) - R_y(\delta)R_x(\delta) = \begin{pmatrix} 1 & -\delta^2 & 0 \\ \delta^2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = R_z(\delta^2) - I, \tag{7.17}$$

where I is the unit matrix. We will show later that this relation can be used to derive the commutation relations between the components of the angular momentum (7.34).

The rotation matrices R are orthogonal:

$$RR^T = R^T R = I, \tag{7.18}$$

where R^T is the transpose of the matrix R. In addition, orthogonal matrices conserve the magnitude of vectors:

$$|\vec{A}'| = |\vec{A}|. \tag{7.19}$$

To see this, since $\vec{A}' = R\vec{A}$, we have $\vec{A}'^2 = (R\vec{A})^2 = \vec{A}^2$ or $A'^2_x + A'^2_y + A'^2_z = A^2_x + A^2_y + A^2_z$.

It is easy to show that the matrices of orthogonal rotations form a (nonabelian) group and that their determinants are equal to 1:

$$\det(R) = 1. \tag{7.20}$$

This group is called the *special* three-dimensional orthogonal group, $SO(3)$, because the rotation group is a special case of a more general group, the group of three-dimensional orthogonal transformations, $O(3)$, which consist of both rotations and reflections and for which we have

$$\det(R) = \pm 1. \tag{7.21}$$

The group $SO(3)$ transforms a vector $\vec{A}$ into another vector $\vec{A}'$ while conserving its magnitude.

7.2 Rotations in Quantum Mechanics

In this section we study the relationship between the angular momentum and the rotation operator and then study the properties as well as the representation of the rotation operator. The connection is analogous to that between the linear momentum operator and translations. We will see that the angular momentum operator acts as a generator for rotations.

A rotation is specified by an angle ϕ and by a unit vector $\vec{n}$ about which the rotation is performed: $\hat{R}_n(\phi)$. Knowing the rotation operator $\hat{R}_n(\phi)$, we can determine how state vectors and operators transform under rotations:

$$|\psi'\rangle = \hat{R}_n |\psi\rangle. \tag{7.22}$$

For the transformation of operators under rotations, we simply need to compare the expectation values $\langle\psi \mid \hat{A} \mid \psi\rangle$ and $\langle\psi' \mid \hat{A}' \mid \psi'\rangle$, where $\hat{A}'$ is the transformed operator of $\hat{A}$ under $\hat{R}_n$. By analogy to the invariance of vectors' magnitudes in classical rotations, $|\vec{A}'| = |\vec{A}|$ (see (7.10)), *the expectation values of operators are also invariant under rotations in quantum mechanics;* to ensure this property, we require that

$$\boxed{\langle\psi' \mid \hat{A}' \mid \psi'\rangle = \langle\psi \mid \hat{A} \mid \psi\rangle \Longrightarrow \langle\psi \mid \hat{R}_n^\dagger \hat{A}' \hat{R}_n \mid \psi\rangle = \langle\psi \mid \hat{A} \mid \psi\rangle \Longrightarrow \hat{A}' = \hat{R}_n \hat{A} \hat{R}_n^\dagger,} \tag{7.23}$$

where we have used the fact that $\hat{R}_n^\dagger \hat{R}_n = \hat{R}_n \hat{R}_n^\dagger = \hat{I}$ as shown in (7.49). The last relation is compatible with the result we have derived in Chapter 2 (see Eq. (2.273)) pertaining to the transformation of an operator under the change of the basis.

After determining how states and operators transform under rotations, we now need to find the rotation operator $\hat{R}_n$. In what follows, we are going to find $\hat{R}_n$ for infinitesimal and for finite rotations.

7.2.1 Infinitesimal Rotations

Consider a rotation of the coordinates of a *spinless* particle over an *infinitesimal* angle $\delta\phi$ about the z-axis. Denoting this rotation by the operator $\hat{R}_z(\delta\phi)$, we have

$$\hat{R}_z(\delta\phi)\, \psi(r, \theta, \phi) = \psi(r, \theta, \phi - \delta\phi). \tag{7.24}$$

Taylor expanding the wave function to the first order in $\delta\phi$, we obtain

$$\psi(r, \theta, \phi - \delta\phi) \simeq \psi(r, \theta, \phi) - \delta\phi \frac{\partial \psi}{\partial \phi} = \left(1 - \delta\phi \frac{\partial}{\partial \phi}\right) \psi(r, \theta, \phi). \tag{7.25}$$

Comparing (7.24) and (7.25) we see that $\hat{R}_z(\delta\phi)$ is given by

$$\hat{R}_z(\delta\phi) = 1 - \delta\phi \frac{\partial}{\partial \phi}. \tag{7.26}$$

Since the z-component of the orbital angular momentum is

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi}, \tag{7.27}$$

we can rewrite (7.26) as

$$\boxed{\hat{R}_z(\delta\phi) = 1 - \frac{i}{\hbar} \delta\phi \hat{L}_z.} \tag{7.28}$$

We may generalize this relation to a rotation of angle $\delta\phi$ about an arbitrary axis whose direction is given by the unit vector $\vec{n}$:

$$\boxed{\hat{R}_n(\delta\phi) = 1 - \frac{i}{\hbar} \delta\phi\, \vec{n} \cdot \hat{\vec{L}}.} \tag{7.29}$$

This is the operator corresponding to an *infinitesimal* rotation of angle $\delta\phi$ about $\vec{n}$ for a *spinless* system. The orbital angular momentum is thus the *generator* of infinitesimal spatial rotations.

Rotations and the commutation relations

We can show that the relation (7.17) leads to the commutation relations of angular momentum $[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z$. The operators corresponding to infinitesimal rotations of angle δ about the x and y axes can be inferred from (7.28):

$$\hat{R}_x(\delta) = 1 - \frac{i\delta}{\hbar}\hat{L}_x - \frac{\delta^2}{2\hbar^2}\hat{L}_x^2, \qquad \hat{R}_y(\delta) = 1 - \frac{i\delta}{\hbar}\hat{L}_y - \frac{\delta^2}{2\hbar^2}\hat{L}_y^2, \tag{7.30}$$

where we have extended the expansions to the second power in δ. On the one hand, the following useful relation can be obtained from (7.30):

$$\begin{aligned}\hat{R}_x(\delta)\hat{R}_y(\delta) - \hat{R}_y(\delta)\hat{R}_x(\delta) &= \left(1 - \frac{i\delta}{\hbar}\hat{L}_x - \frac{\delta^2}{2\hbar^2}\hat{L}_x^2\right)\left(1 - \frac{i\delta}{\hbar}\hat{L}_y - \frac{\delta^2}{2\hbar^2}\hat{L}_y^2\right) \\ &\quad - \left(1 - \frac{i\delta}{\hbar}\hat{L}_y - \frac{\delta^2}{2\hbar^2}\hat{L}_y^2\right)\left(1 - \frac{i\delta}{\hbar}\hat{L}_x - \frac{\delta^2}{2\hbar^2}\hat{L}_x^2\right) \\ &= -\frac{\delta^2}{\hbar^2}\left(\hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x\right) \\ &= -\frac{\delta^2}{\hbar^2}[\hat{L}_x, \hat{L}_y], \end{aligned} \tag{7.31}$$

where we have kept only terms up to the second power in δ; the terms in δ cancel out automatically.

On the other hand, according to (7.17), we have

$$R_x(\delta)R_y(\delta) - R_y(\delta)R_x(\delta) = R_z(\delta^2) - 1. \tag{7.32}$$

Since $\hat{R}_z(\delta^2) = 1 - (i\delta^2/\hbar)\hat{L}_z$ this relations leads to

$$R_x(\delta)R_y(\delta) - R_y(\delta)R_x(\delta) = R_z(\delta^2) - 1 = -\frac{i\delta^2}{\hbar}\hat{L}_z. \tag{7.33}$$

Finally, equating (7.31) and (7.33), we end up with

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z. \tag{7.34}$$

Similar calculations for $R_y(\delta)R_z(\delta) - R_z(\delta)R_y(\delta)$ and $R_z(\delta)R_x(\delta) - R_x(\delta)R_z(\delta)$ lead to the other two commutation relations $[\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x$ and $[\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y$.

7.2.2 Finite Rotations

The operator $\hat{R}_z(\phi)$ corresponding to a rotation (of the coordinates of a spinless particle) over a *finite* angle ϕ about the z-axis can be constructed in terms of the infinitesimal rotation operator (7.28) as follows. We divide the angle ϕ into N infinitesimal angles $\delta\phi$: $\phi = N\,\delta\phi$. The rotation over the finite angle ϕ can thus be viewed as a series of N consecutive infinitesimal rotations, each over the angle $\delta\phi$, about the z-axis, applied consecutively one after the other:

$$\hat{R}_z(\phi) = \hat{R}_z(N\delta\phi) = R_z(\delta\phi)R_z(\delta\phi)\cdots R_z(\delta\phi) = \left(R_z(\delta\phi)\right)^N = \left(1 - i\frac{\delta\phi}{\hbar}\hat{L}_z\right)^N. \tag{7.35}$$

Since $\delta\phi = \phi/N$, and if $\delta\phi$ is infinitesimally small, we have

$$\hat{R}_z(\phi) = \lim_{N\to\infty} \prod_{k=1}^{N} \left(1 - \frac{i}{\hbar}\frac{\phi}{N}\vec{n}\cdot\hat{\vec{L}}\right) = \lim_{N\to\infty}\left(1 - \frac{i}{\hbar}\frac{\phi}{N}\hat{L}_z\right)^N, \tag{7.36}$$

which, when combined with the well-known expression $\lim_{N\to\infty}(1 + x/N)^N = e^x$, leads to

$$\boxed{\hat{R}_z(\phi) = e^{-i\phi\hat{L}_z/\hbar}.} \tag{7.37}$$

We can generalize this result to infer the rotation operator $\hat{R}_n(\phi)$ corresponding to a rotation over a finite angle ϕ around an axis $\vec{n}$:

$$\boxed{\hat{R}_n(\phi) = e^{-i\phi\vec{n}\cdot\hat{\vec{L}}/\hbar},} \tag{7.38}$$

where $\hat{\vec{L}}$ is the orbital angular momentum. This operator represents the rotation of the coordinates of a spinless particle over an angle ϕ about an axis $\vec{n}$.

The discussion that led to (7.38) was carried out for a spinless system. A more general study for a system with spin would lead to a relation similar to (7.38):

$$\boxed{\hat{R}_n(\phi) = e^{-\frac{i}{\hbar}\phi\vec{n}\cdot\hat{\vec{J}}},} \tag{7.39}$$

where $\hat{\vec{J}}$ is the total angular momentum operator; this is known as the *rotation operator*. For instance, the rotation operator $\hat{R}_x(\phi)$ of a rotation through an angle ϕ about the x-axis is given by

$$\hat{R}_x(\phi) = e^{-i\phi\hat{J}_x/\hbar}. \tag{7.40}$$

The properties of $\hat{R}_n(\phi)$ are determined by those of the operators $\hat{J}_x, \hat{J}_y, \hat{J}_z$.

Remark

The Hamiltonian of a particle in a *central potential*, $\hat{H} = \hat{P}^2/(2m) + \hat{V}(r)$, is *invariant under spatial rotations* since, as shown in Chapter 6, it commutes with the orbital angular momentum:

$$[\hat{H}, \hat{\vec{L}}] = 0 \quad \Longrightarrow \quad \left[\hat{H}, e^{-i\phi\vec{n}\cdot\hat{\vec{L}}/\hbar}\right] = 0. \tag{7.41}$$

Due to this symmetry of space isotropy or rotational invariance, the *orbital angular momentum is conserved*[1]. So, in the case of particles moving in central potentials, the orbital angular momentum is a constant of the motion.

7.2.3 Properties of the Rotation Operator

The rotation operators constitute a representation of the rotation group and satisfy the following properties:

[1] In classical physics when a system is invariant under rotations, its total angular momentum is also conserved.

- The product of any two rotation operators is another rotation operator:

$$\hat{R}_{n_1}\hat{R}_{n_2} = \hat{R}_{n_3}. \tag{7.42}$$

- The associative law holds for rotation operators:

$$\left(\hat{R}_{n_1}\hat{R}_{n_2}\right)\hat{R}_{n_3} = \hat{R}_{n_1}\left(\hat{R}_{n_2}\hat{R}_{n_3}\right). \tag{7.43}$$

- The identity operator (corresponding to no rotation) satisfies the relation

$$\hat{I}\hat{R}_n = \hat{R}_n\hat{I} = \hat{R}_n. \tag{7.44}$$

From (7.39) we see that for each rotation operator $\hat{R}_n$, there exists an inverse operator $\hat{R}_n^{-1}$ so that

$$\hat{R}_n\hat{R}_n^{-1} = \hat{R}_n^{-1}\hat{R}_n = \hat{I}. \tag{7.45}$$

The operator $\hat{R}_{-n}$, which is equal to $\hat{R}_n^{-1}$, corresponds to a rotation in the opposite sense to $\hat{R}_n$.

In sharp contrast to the translation group[2] in three dimensions, the rotation group is not commutative (nonabelian). *The product of two rotation operators depends on the order in which they are performed*:

$$\hat{R}_{n_1}(\phi)\hat{R}_{n_2}(\theta) \neq \hat{R}_{n_2}(\theta)\hat{R}_{n_1}(\phi); \tag{7.46}$$

this is due to the fact that the commutator $[\vec{n}_1 \cdot \hat{\vec{J}},\ \vec{n}_2 \cdot \hat{\vec{J}}]$ is not zero. In this way, the rotation group is in general nonabelian.

But if the two rotations were performed about the *same* axis, the corresponding operators would *commute*:

$$\hat{R}_n(\phi)\hat{R}_n(\theta) = \hat{R}_n(\theta)\hat{R}_n(\phi) = \hat{R}_n(\phi + \theta). \tag{7.47}$$

Note that, since the angular momentum operator $\hat{J}$ is Hermitian, equation (7.39) yields

$$\boxed{\hat{R}_n^\dagger(\phi) = \hat{R}_n^{-1}(\phi) = \hat{R}_n(-\phi) = e^{i\phi\vec{n}\cdot\hat{\vec{J}}/\hbar};} \tag{7.48}$$

hence the rotation operator (7.39) is unitary:

$$\boxed{\hat{R}_n^\dagger(\phi) = \hat{R}_n^{-1}(\phi) \quad \Longrightarrow \quad \hat{R}_n^\dagger(\phi)\hat{R}_n(\phi) = \hat{I}.} \tag{7.49}$$

Hence, the operator $\hat{R}_n(\phi)$ conserves the scalar product of kets, notably the norm of vectors. For instance, using

$$|\,\psi'\rangle = \hat{R}_n(\phi)\,|\,\psi\rangle, \qquad |\,\chi'\rangle = \hat{R}_n(\phi)\,|\,\chi\rangle, \tag{7.50}$$

along with (7.49), we can show that $\langle\chi' \mid \psi'\rangle = \langle\chi \mid \psi\rangle$:

$$\langle\chi' \mid \psi'\rangle = \langle\chi \mid \hat{R}_n^\dagger(\phi)\hat{R}_n(\phi) \mid \psi\rangle = \langle\chi \mid \psi\rangle. \tag{7.51}$$

[2]The linear momenta $\hat{P}_i$ and $\hat{P}_j$—which are the generators of translation—commute even when $i \neq j$; hence the translation group is said to be abelian.

7.2.4 Euler Rotations

From classical mechanics, we know that an arbitrary rotation of a rigid body can be expressed in terms of three consecutive rotations, called the Euler rotations. In quantum mechanics, instead of expressing the rotation operator $\hat{R}_{\vec{n}}(\phi) = e^{-i\phi\vec{n}\cdot\vec{J}/\hbar}$ in terms of a rotation through an angle ϕ about an arbitrary axis $\vec{n}$, it is more convenient to parameterize it, as in classical mechanics, in terms of the three *Euler angles* (α, β, γ) where $0 \leq \alpha \leq 2\pi$, $0 \leq \beta \leq \pi$, and $0 \leq \gamma \leq 2\pi$. The Euler rotations transform the space-fixed set of axes xyz into a new set $x'y'z'$, having the same origin O, by means of three consecutive counterclockwise rotations:

- First, rotate the space-fixed $Oxyz$ system through an angle α about the z-axis; this rotation transforms the $Oxyz$ system into $Ouvz$: $Oxyz \longrightarrow Ouvz$.
- Second, rotate the uvz system through an angle β about the v-axis; this rotation transforms the $Ouvz$ system into $Owvz'$: $Ouvz \longrightarrow Owvz'$.
- Third, rotate the wvz' system through an angle γ about the z'-axis; this rotation transforms the $Owvz'$ system into $Ox'y'z'$: $Owvz' \longrightarrow Ox'y'z'$.

The operators representing these three rotations are given by $\hat{R}_z(\alpha)$, $\hat{R}_v(\beta)$, and $\hat{R}_{z'}(\gamma)$, respectively. Using (7.39) we can represent these three rotations by

$$\hat{R}(\alpha,\beta,\gamma) = \hat{R}_{z'}(\gamma)\hat{R}_v(\beta)\hat{R}_z(\alpha) = \exp\left[-i\gamma J_{z'}/\hbar\right]\exp\left[-i\beta J_v/\hbar\right]\exp\left[-i\alpha J_z/\hbar\right]. \tag{7.52}$$

The form of this operator is rather inconvenient, for it includes rotations about axes belonging to different systems (i.e., z', v, and z); this form would be most convenient were we to express (7.52) as a product of three rotations about the space-fixed axes x, y, z. So let us express $\hat{R}_{z'}(\gamma)$ and $\hat{R}_v(\beta)$ in terms of rotations about the x, y, z axes. Since the first Euler rotation described above, $\hat{R}_z(\alpha)$, transforms the operator $\hat{J}_y$ into $\hat{J}_v$, i.e., $\hat{J}_v = \hat{R}_z(\alpha)\hat{J}_y\hat{R}_z(-\alpha)$ by (7.23), we have

$$\hat{R}_v(\beta) = \hat{R}_z(\alpha)\hat{R}_y(\beta)\hat{R}_z(-\alpha) = e^{-i\alpha J_z/\hbar}e^{-i\beta J_y/\hbar}e^{i\alpha J_z/\hbar}. \tag{7.53}$$

Here $\hat{J}_{z'}$ is obtained from $\hat{J}_z$ by the consecutive application of the second and third Euler rotations, $\hat{J}_{z'} = \hat{R}_v(\beta)\hat{R}_z(\alpha)\hat{J}_z\hat{R}_z(-\alpha)\hat{R}_v(-\beta)$; hence

$$\hat{R}_{z'}(\gamma') = \hat{R}_v(\beta)\hat{R}_z(\alpha)\hat{R}_z(\gamma)\hat{R}_z(-\alpha)\hat{R}_v(-\beta). \tag{7.54}$$

Since $\hat{R}_v(-\beta) = \hat{R}_z(\alpha)\hat{R}_y(-\beta)\hat{R}_z(-\alpha)$, substituting (7.53) into (7.54) we obtain

$$\begin{aligned}\hat{R}_{z'}(\gamma') &= \left[\hat{R}_z(\alpha)\hat{R}_y(\beta)\hat{R}_z(-\alpha)\right]\hat{R}_z(\alpha)\hat{R}_z(\gamma)\hat{R}_z(-\alpha)\left[\hat{R}_z(\alpha)\hat{R}_y(-\beta)\hat{R}_z(-\alpha)\right]\\ &= \hat{R}_z(\alpha)\hat{R}_y(\beta)\hat{R}_z(\gamma)\hat{R}_y(-\beta)\hat{R}_z(-\alpha)\\ &= e^{-i\alpha J_z/\hbar}e^{-i\beta J_y/\hbar}e^{-i\gamma' J_z/\hbar}e^{i\beta J_y/\hbar}e^{i\alpha J_z/\hbar},\end{aligned} \tag{7.55}$$

where we used the fact that $\hat{R}_z(-\alpha)\hat{R}_z(\alpha) = e^{-i\alpha J_z/\hbar}e^{i\alpha J_z/\hbar} = 1$.

Finally, inserting (7.53) and (7.55) into (7.52) and simplifying (i.e., using $\hat{R}_z(-\alpha)\hat{R}_z(\alpha) = 1$ and $\hat{R}_y(-\beta)\hat{R}_y(\beta) = 1$), we end up with a product of three rotations about the space-fixed axes y and z:

$$\boxed{\hat{R}(\alpha,\beta,\gamma) = \hat{R}_z(\alpha)\hat{R}_y(\beta)\hat{R}_z(\gamma) = e^{-i\alpha J_z/\hbar}e^{-i\beta J_y/\hbar}e^{-i\gamma J_z/\hbar}.} \tag{7.56}$$

The inverse transformation of (7.56) is obtained by taking three rotations in reverse order over the angles $(-\gamma, -\beta, -\alpha)$:

$$\boxed{\hat{R}^{-1}(\alpha,\beta,\gamma) = \hat{R}_z(-\gamma)\hat{R}_y(-\beta)\hat{R}_z(-\alpha) = \hat{R}^{\dagger}(\alpha,\beta,\gamma) = e^{i\gamma J_z/\hbar}e^{i\beta J_y/\hbar}e^{i\alpha J_z/\hbar}.} \tag{7.57}$$

7.2.5 Representation of the Rotation Operator

The rotation operator $\hat{R}(\alpha,\beta,\gamma)$ as given by (7.56) implies that its properties are determined by the algebraic properties of the angular momentum operators $\hat{J}_x$, $\hat{J}_y$, $\hat{J}_z$. Since $\hat{R}(\alpha,\beta,\gamma)$ commutes with $\hat{\vec{J}}^2$, we may look for a representation of $\hat{R}(\alpha,\beta,\gamma)$ in the basis spanned by the eigenvectors of $\hat{\vec{J}}^2$ and J_z, i.e., the $|\, j,\ m\rangle$ states.

From (7.56), we see that $\hat{\vec{J}}^2$ commutes with the rotation operator, $[\hat{\vec{J}}^2, \hat{R}(\alpha,\beta,\gamma)] = 0$; thus, the total angular momentum is conserved under rotations

$$\hat{\vec{J}}^2\hat{R}(\alpha,\beta,\gamma)\,|\, j,\ m\rangle = \hat{R}(\alpha,\beta,\gamma)\hat{\vec{J}}^2\,|\, j,\ m\rangle = \hbar^2 j(j+1)\hat{R}(\alpha,\beta,\gamma)\,|\, j,\ m\rangle. \tag{7.58}$$

However, the z-component of the angular momentum changes under rotations, unless the axis of rotation is along the z-axis. That is, when $\hat{R}(\alpha,\beta,\gamma)$ acts on the state $|\, j,\ m\rangle$, we end up with a new state having the same j but with a different value of m:

$$\begin{aligned}\hat{R}(\alpha,\beta,\gamma)\,|\, j,\ m\rangle &= \sum_{m'=-j}^{j} |\, j,\ m'\rangle\langle j,\ m'\,|\,\hat{R}(\alpha,\beta,\gamma)\,|\, j,\ m\rangle \\ &= \sum_{m'=-j}^{j} D^{(j)}_{m'm}(\alpha,\beta,\gamma)\,|\, j,\ m'\rangle,\end{aligned} \tag{7.59}$$

where

$$\boxed{D^{(j)}_{m'm}(\alpha,\beta,\gamma) = \langle j,\ m'\,|\,\hat{R}(\alpha,\beta,\gamma)\,|\, j,\ m\rangle.} \tag{7.60}$$

These are the matrix elements of $\hat{R}(\alpha,\beta,\gamma)$ for the $|\, j,\ m\rangle$ states; $D^{(j)}_{m'm}(\alpha,\beta,\gamma)$ is the amplitude of $|\, j,\ m'\rangle$ when $|\, j,\ m\rangle$ is rotated. The rotation operator is thus represented by a $(2j+1)\times(2j+1)$ square matrix in the $\{|\, j,\ m\rangle\}$ basis. The matrix of $D^{(j)}(\alpha,\beta,\gamma)$ is known as the *Wigner D-matrix* and its elements $D^{(j)}_{m'm}(\alpha,\beta,\gamma)$ as the *Wigner functions*. This matrix representation is often referred to as the $(2j+1)$-dimensional *irreducible representation* of the rotation operator $\hat{R}(\alpha,\beta,\gamma)$.

Since $|\, j,\ m\rangle$ is an eigenstate of J_z, it is also an eigenstate of the rotation operator $e^{i\alpha J_z/\hbar}$, because

$$e^{i\alpha J_z/\hbar}\,|\, j,\ m\rangle = e^{i\alpha m}\,|\, j,\ m\rangle. \tag{7.61}$$

We may thus rewrite (7.60) as

$$\boxed{D^{(j)}_{m'm}(\alpha,\beta,\gamma) = e^{-i(m'\alpha+m\gamma)}d^{(j)}_{m'm}(\beta),} \tag{7.62}$$

where

$$\boxed{d^{(j)}_{m'm}(\beta) = \langle j,\ m'|e^{-i\beta J_y/\hbar} \mid j,\ m\rangle.} \tag{7.63}$$

This shows that only the middle rotation operator, $e^{-i\beta J_y/\hbar}$, mixes states with different values of m. Determining the matrix elements $D^{(j)}_{m'm}(\alpha,\beta,\gamma)$ therefore reduces to evaluation of the quantities $d^{(j)}_{m'm}(\beta)$.

A general expression of $d^{(j)}_{m'm}(\beta)$, called the *Wigner formula*, is given by the following explicit expression:

$$\begin{aligned} d^{(j)}_{m'm}(\beta) &= \sum_k (-1)^{k+m'-m} \frac{\sqrt{(j+m)!(j-m)!(j+m')!(j-m')!}}{(j-m'-k)!(j+m-k)!(k+m'-m)!k!} \\ &\quad \times \left(\cos\frac{\beta}{2}\right)^{2j+m-m'-2k} \left(\sin\frac{\beta}{2}\right)^{m'-m+2k}. \end{aligned} \tag{7.64}$$

The summation over k is taken such that none of the arguments of factorials in the denominator are negative.

We should note that, since the D-function $D^{(j)}_{m'm}(\alpha,\beta,\gamma)$ is a joint eigenfunction of $\hat{J}^2$ and J_z, we have

$$\boxed{\hat{J}^2 D^{(j)}_{m'm}(\alpha,\beta,\gamma) = j(j+1)\hbar^2 D^{(j)}_{m'm}(\alpha,\beta,\gamma),} \tag{7.65}$$

$$\boxed{\hat{J}_z D^{(j)}_{m'm}(\alpha,\beta,\gamma) = \hbar m D^{(j)}_{m'm}(\alpha,\beta,\gamma),} \tag{7.66}$$

$$\boxed{\hat{J}_\pm D^{(j)}_{m'm}(\alpha,\beta,\gamma) = \hbar\sqrt{(j\pm m)(j\mp m+1)}\, D^{(j)}_{m'm\pm1}(\alpha,\beta,\gamma).} \tag{7.67}$$

Properties of the D-functions

We now list some of the most useful properties of the rotation matrices. The complex conjugate of the D-functions can be expressed as

$$\begin{aligned} \left[D^{(j)}_{m'm}(\alpha,\beta,\gamma)\right]^* &= \langle j,\ m' \mid \hat{R}(\alpha,\beta,\gamma) \mid j,\ m\rangle^* = \langle j\, m \mid \hat{R}^\dagger(\alpha,\beta,\gamma) \mid j,\ m'\rangle \\ &= \langle j\, m \mid \hat{R}^{-1}(\alpha,\beta,\gamma) \mid j,\ m'\rangle \\ &= D^{(j)}_{mm'}(-\gamma,-\beta,-\alpha). \end{aligned} \tag{7.68}$$

We can easily show that

$$\boxed{\left[D^{(j)}_{m'm}(\alpha,\beta,\gamma)\right]^* = (-1)^{m'-m} D^{(j)}_{-m'-m}(\alpha,\beta,\gamma) = D^{(j)}_{mm'}(-\gamma,-\beta,-\alpha).} \tag{7.69}$$

The D-functions satisfy the following unitary relations:

$$\boxed{\sum_m \left[D^{(j)}_{km}(\alpha,\beta,\gamma)\right]^* D^{(j)}_{k'm}(\alpha,\beta,\gamma) = \delta_{k,k'},} \tag{7.70}$$

$$\boxed{\sum_m \left[D^{(j)}_{mk}(\alpha,\beta,\gamma)\right]^* D^{(j)}_{mk'}(\alpha,\beta,\gamma) = \delta_{k,k'},} \tag{7.71}$$

since

$$\sum_m \left[D^{(j)}_{mk}(\alpha,\beta,\gamma)\right]^* D^{(j)}_{mk'}(\alpha,\beta,\gamma) = \sum_m \langle j\,k|\hat{R}^{-1}(\alpha,\beta,\gamma) \mid j,\ m\rangle\langle j\,m \mid R(\alpha,\beta,\gamma)|j\,k'\rangle$$
$$= \langle j\,k|\hat{R}^{-1}(\alpha,\beta,\gamma)\hat{R}(\alpha,\beta,\gamma)(|j\,k'\rangle$$
$$= \langle j\,k|j\,k'\rangle$$
$$= \delta_{k,k'}. \tag{7.72}$$

From (7.63) we can show that the d-functions satisfy the following relations:

$$d^{(j)}_{m'm}(\pi) = (-1)^{j-m}\delta_{m',-m}, \qquad d^{(j)}_{m'm}(0) = \delta_{m',m}. \tag{7.73}$$

Since $d^j_{m'm}$ are elements of a unitary real matrix, the matrix $d^{(j)}(\beta)$ must be orthogonal. We may thus write

$$\boxed{d^{(j)}_{m'm}(\beta) = \left(d^{(j)}_{m'm}(\beta)\right)^{-1} = d^{(j)}_{mm'}(-\beta)} \tag{7.74}$$

and

$$\boxed{d^{(j)}_{m'm}(\beta) = (-1)^{m'-m}d^{(j)}_{mm'}(\beta) = (-1)^{m'-m}d^{(j)}_{-m'-m}(\beta).} \tag{7.75}$$

The unitary matrices $D^{(j)}$ form a $(2j+1)$ dimensional irreducible representation of the $SO(3)$ group.

7.2.6 Rotation Matrices and the Spherical Harmonics

In the case where the angular momentum operator $\vec{J}$ is purely orbital (i.e., the values of j are integer, $j = l$), there exists a connection between the D-functions and the spherical harmonics $Y_{lm}(\theta,\varphi)$. The operator $\hat{R}(\alpha,\beta,\gamma)$ when applied to a vector $|\vec{r}\rangle$ pointing in the direction (θ,φ) would generate a vector $|\vec{r}'\rangle$ along a new direction (θ',φ'):

$$|\vec{r}'\rangle = \hat{R}(\alpha,\beta,\gamma)\,|\vec{r}\rangle. \tag{7.76}$$

An expansion in terms of $|l,\ m'\rangle$ and a multiplication by $\langle l,\ m|$ leads to

$$\langle l,\ m \mid \vec{r}'\rangle = \sum_{m'}\langle l,\ m|\hat{R}(\alpha,\beta,\gamma) \mid l,\ m'\rangle\langle l,\ m' \mid \vec{r}\rangle, \tag{7.77}$$

or to

$$Y^*_{lm}(\theta',\varphi') = \sum_{m'} D^{(l)}_{m\,m'}(\alpha,\beta,\gamma)Y^*_{lm'}(\theta,\varphi), \tag{7.78}$$

since $\langle l,\ m \mid \vec{r}'\rangle = Y^*_{lm}(\theta',\varphi')$ and $\langle l,\ m' \mid \vec{r}\rangle = Y^*_{lm'}(\theta,\varphi)$.

In the case where the vector $\vec{r}$ is along the z-axis, we have $\theta = 0$; hence $m' = 0$. From Chapter 5, $Y^*_{l0}(0,\varphi)$ is given by

$$Y^*_{lm'}(0,\varphi) = \sqrt{\frac{2l+1}{4\pi}}\delta_{m',0}. \tag{7.79}$$

We can thus reduce (7.78) to

$$Y^*_{lm}(\beta,\ \alpha) = D^{(l)}_{m\ 0}(\alpha,\beta,\gamma)Y^*_{l0}(0,\varphi) = \sqrt{\frac{2l+1}{4\pi}}D^{(l)}_{m\ 0}(\alpha,\beta,\gamma), \tag{7.80}$$

or to

$$\boxed{D^{(l)}_{m0}(\alpha,\beta,\gamma) = \sqrt{\frac{4\pi}{2l+1}}Y^*_{lm}(\beta,\ \alpha).} \tag{7.81}$$

This means that a rotation through the Euler angles (α,β,γ) of the vector $\vec{r}$, when it is along the z-axis, produces a vector $\vec{r}\,'$ whose azimuthal and polar angles are given by β and α, respectively. Similarly, we can show that

$$\boxed{D^{(l)}_{0m}(\gamma,\beta,\alpha) = \sqrt{\frac{4\pi}{2l+1}}Y_{lm}(\beta,\ \alpha)} \tag{7.82}$$

and

$$D^{(l)}_{00}(0,\theta,0) = P_l(\cos\ \theta), \tag{7.83}$$

where $P_l(\cos\ \theta)$ is the Legendre polynomial.

We are now well equipped to derive the theorem for the addition of spherical harmonics. Let (θ,φ) be the polar coordinates of the vector $\vec{r}$ with respect to the space-fixed x,y,z system and let (θ',φ') be its polar coordinates with respect to the rotated system x',y',z'; taking the complex conjugate of (7.78) we obtain

$$Y_{lm}(\theta',\varphi') = \sum_{m'}\left[D^{(l)}_{m\ m'}(\alpha,\beta,\gamma)\right]^*Y_{lm'}(\theta,\varphi). \tag{7.84}$$

For the case $m = 0$, since (from Chapter 5)

$$Y_{l0}(\theta',\varphi') = \sqrt{\frac{2l+1}{4\pi}}P_l(\cos\ \theta') \tag{7.85}$$

and since from (7.82)

$$\left[D^{(l)}_{0m'}(\alpha,\beta,\gamma)\right]^* = \sqrt{\frac{4\pi}{2l+1}}Y^*_{lm'}(\beta,\gamma), \tag{7.86}$$

we can reduce (7.84) to

$$\sqrt{\frac{2l+1}{4\pi}}P_l(\cos\ \theta') = \sum_{m'}\sqrt{\frac{4\pi}{2l+1}}Y^*_{lm'}(\beta,\gamma)Y_{lm'}(\theta,\varphi), \tag{7.87}$$

or to

$$P_l(\cos\ \theta') = \frac{4\pi}{2l+1}\sum_{m'}Y^*_{lm'}(\beta,\gamma)Y_{lm'}(\theta,\varphi). \tag{7.88}$$

Integrals involving the D-functions

Let ω denote the Euler angles; hence

$$\int d\omega = \int_0^{\pi} \sin\beta d\beta \int_0^{2\pi} d\alpha \int_0^{2\pi} d\gamma. \tag{7.89}$$

Using the relation

$$\begin{aligned} \int D^{(j)}_{m'm}(\omega)\, d\omega &= \int_0^{\pi} d^{(j)}_{m'm}(\beta) \sin\beta d\beta \int_0^{2\pi} e^{-im'\alpha} d\alpha \int_0^{2\pi} e^{-im\gamma'} d\gamma' \\ &= 8\pi^2 \delta_{j,0}\delta_{m',0}\delta_{m,0}, \end{aligned} \tag{7.90}$$

we may write

$$\begin{aligned} \int D^{(j)*}_{mk}(\omega) D^{(j')}_{m'k'}(\omega)\, d\omega &= (-1)^{m-k} \int D^{(j)}_{-m-k}(\omega) D^{(j')}_{m'k'}(\omega)\, d\omega \\ &= (-1)^{m-k} \int_0^{\pi} d^{(j)}_{-m-k}(\beta) d^{(j')}_{m'k'}(\beta) \sin\beta d\beta \\ &\quad \times \int_0^{2\pi} e^{-i(m'-m)\alpha} d\alpha \int_0^{2\pi} e^{-i(k'-k)\gamma'} d\gamma' \\ &= \frac{8\pi^2}{2j+1} \delta_{j,j'}\delta_{m,m'}\delta_{k,k'}. \end{aligned} \tag{7.91}$$

Example 7.1
Find the rotation matrices $d^{(1/2)}$ and $D^{(1/2)}$ corresponding to $j = \frac{1}{2}$.

Solution
On the one hand, since the matrix of $\hat{J}_y$ for $j = \frac{1}{2}$ (Chapter 5) is given by

$$\hat{J}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{\hbar}{2}\sigma_y, \tag{7.92}$$

and since the square of the Pauli matrix σ_y is equal to the unit matrix, $\sigma_y^2 = 1$, the even and odd powers of σ_y are given by

$$\sigma_y^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \sigma_y^{2n+1} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_y. \tag{7.93}$$

On the other hand, since the rotation operator

$$\hat{R}_y(\beta) = e^{-i\beta\hat{J}_y/\hbar} = e^{-i\beta\sigma_y/2} \tag{7.94}$$

can be written as

$$e^{-i\beta\sigma_y/2} = \sum_{n=0}^{\infty} \frac{(-i)^{2n}}{(2n)!} \left(\frac{\beta}{2}\right)^{2n} \sigma^{2n} + \sum_{n=0}^{\infty} \frac{(-i)^{2n+1}}{(2n+1)!} \left(\frac{\beta}{2}\right)^{2n+1} \sigma_y^{2n+1}, \tag{7.95}$$

a substitution of (7.93) into (7.95) yields

$$e^{-i\beta\sigma_y/2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \left(\frac{\beta}{2}\right)^{2n} - i\sigma_y \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \left(\frac{\beta}{2}\right)^{2n+1}$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cos\left(\frac{\beta}{2}\right) + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \sin\left(\frac{\beta}{2}\right); \tag{7.96}$$

hence

$$d^{(1/2)}(\beta) = e^{-i\beta J_y/\hbar} = \begin{pmatrix} d^{(1/2)}_{\frac{1}{2}\frac{1}{2}} & d^{(1/2)}_{\frac{1}{2}-\frac{1}{2}} \\ d^{(1/2)}_{-\frac{1}{2}\frac{1}{2}} & d^{(1/2)}_{-\frac{1}{2}-\frac{1}{2}} \end{pmatrix} = \begin{pmatrix} \cos(\beta/2) & -\sin(\beta/2) \\ \sin(\beta/2) & \cos(\beta/2) \end{pmatrix}. \tag{7.97}$$

Since as shown in (7.62) $D^{(j)}_{m'm}(\alpha,\beta,\gamma') = e^{-i(m'\alpha+m\gamma)} d^{(j)}_{m'm}(\beta)$, we have

$$D^{(1/2)}(\alpha,\beta,\gamma') = \begin{pmatrix} e^{-i(\alpha+\gamma')/2}\cos(\beta/2) & -e^{-i(\alpha-\gamma')/2}\sin(\beta/2) \\ e^{i(\alpha-\gamma')/2}\sin(\beta/2) & e^{i(\alpha+\gamma')/2}\cos(\beta/2) \end{pmatrix}. \tag{7.98}$$

7.3 Addition of Angular Momenta

The addition of angular momenta is encountered in all areas of modern physics. Mastering its techniques is essential for an understanding of the various subatomic phenomena. For instance, the total angular momentum of the electron in a hydrogen atom consists of two parts, an orbital part $\hat{\vec{L}}$, which is due to the orbiting motion of the electron around the proton, and a spin part $\hat{\vec{S}}$, which is due to the spinning motion of the electron about itself. The properties of the hydrogen atom cannot be properly discussed without knowing how to add the orbital and spin parts of the electron's total angular momentum.

In what follows we are going to present the formalism of angular momentum addition and then consider some of its most essential applications.

7.3.1 Addition of Two Angular Momenta: General Formalism

In this section we present the general formalism corresponding to the problem of adding two *commuting* angular momenta.

Consider two angular momenta $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ which belong to different subspaces 1 and 2; $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ may refer to two distinct particles or to two different properties of the same particle[3]. The latter case may refer to the orbital and spin angular momenta of the same particle.

[3]Throughout this section we shall use the labels 1 and 2 to refer to quantities relevant to the two particles or the two subspaces.

Assuming that the spin–orbit coupling is sufficiently weak, then the space and spin degrees of freedom of the electron evolve *independently* of each other.

The components of $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ satisfy the usual commutation relations of angular momentum:

$$\left[\hat{J}_{1_x}, \hat{J}_{1_y}\right] = i\hbar\hat{J}_{1_z}, \qquad \left[\hat{J}_{1_y}, \hat{J}_{1_z}\right] = i\hbar\hat{J}_{1_x}, \qquad \left[\hat{J}_{1_z}, \hat{J}_{1_x}\right] = i\hbar\hat{J}_{1_y}, \tag{7.99}$$

$$\left[\hat{J}_{2_x}, \hat{J}_{2_y}\right] = i\hbar\hat{J}_{2_z}, \qquad \left[\hat{J}_{2_y}, \hat{J}_{2_z}\right] = i\hbar\hat{J}_{2_x}, \qquad \left[\hat{J}_{2_z}, \hat{J}_{2_x}\right] = i\hbar\hat{J}_{2_y}. \tag{7.100}$$

Since $\hat{\vec{J}}_1$, and $\hat{\vec{J}}_2$ belong to different spaces, their components commute:

$$\left[\hat{J}_{1_j}, \hat{J}_{2_k}\right] = 0, \qquad (j, k = x, y, z). \tag{7.101}$$

Now, denoting the joint eigenstates of $\hat{\vec{J}}_1^2$ and $\hat{J}_{1_z}$ by $|\, j_1, m_1\rangle$ and those of $\hat{\vec{J}}_2^2$ and $\hat{J}_{2_z}$ by $|\, j_2, m_2\rangle$, we have

$$\hat{\vec{J}}_1^2 \,|\, j_1, m_1\rangle = j_1(j_1+1)\hbar^2 \,|\, j_1, m_1\rangle, \tag{7.102}$$

$$\hat{J}_{1_z} \,|\, j_1, m_1\rangle = m_1\hbar \,|\, j_1, m_1\rangle, \tag{7.103}$$

$$\hat{\vec{J}}_2^2 \,|\, j_2, m_2\rangle = j_2(j_2+1)\hbar^2 \,|\, j_2, m_2\rangle, \tag{7.104}$$

$$\hat{J}_{2_z} \,|\, j_2, m_2\rangle = m_2\hbar \,|\, j_2, m_2\rangle. \tag{7.105}$$

The dimensions of the spaces to which $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ belong are given by $(2j_1 + 1)$ and $(2j_2 + 1)$, respectively[4]. The operators $\hat{\vec{J}}_1^2$ and $\hat{J}_{1_z}$ are represented within the $\{|\, j_1, m_1\rangle\}$ basis by square matrices of dimension $(2j_1 + 1) \times (2j_1 + 1)$, while $\hat{\vec{J}}_2^2$ and $\hat{J}_{2_z}$ are represented by square matrices of dimension $(2j_2 + 1) \times (2j_2 + 1)$ within the $\{|\, j_2, m_2\rangle\}$ basis.

Consider now the two particles (or two subspaces) 1 and 2 together. The four operators $\hat{\vec{J}}_1^2$, $\hat{\vec{J}}_2^2$, $\hat{J}_{1_z}$, $\hat{J}_{2_z}$ form a complete set of commuting operators; they can thus be jointly diagonalized by the same states. Denoting their joint eigenstates by $|\, j_1, j_2; m_1, m_2\rangle$, we can write them as *direct products* of $|\, j_1, m_1\rangle$, and $|\, j_2, m_2\rangle$

$$|\, j_1, j_2; m_1, m_2\rangle = |\, j_1, m_1\rangle \,|\, j_2, m_2\rangle, \tag{7.106}$$

because the coordinates of $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ are *independent*. We can thus rewrite (7.102)–(7.105) as

$$\hat{\vec{J}}_1^2 \,|\, j_1, j_2; m_1, m_2\rangle = j_1(j_1+1)\hbar^2 \,|\, j_1, j_2; m_1, m_2\rangle, \tag{7.107}$$

$$\hat{J}_{1_z} \,|\, j_1, j_2; m_1, m_2\rangle = m_1\hbar \,|\, j_1, j_2; m_1, m_2\rangle, \tag{7.108}$$

$$\hat{\vec{J}}_2^2 \,|\, j_1, j_2; m_1, m_2\rangle = j_2(j_2+1)\hbar^2 \,|\, j_1, j_2; m_1, m_2\rangle, \tag{7.109}$$

$$\hat{J}_{2_z} \,|\, j_1, j_2; m_1, m_2\rangle = m_2\hbar \,|\, j_1, j_2; m_1, m_2\rangle. \tag{7.110}$$

[4]This is due to the fact that the number of basis vectors spanning the spaces to which $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ belong are equal to $(2j_1 + 1)$ and $(2j_2 + 1)$, respectively; these vectors are $|j_1, -j_1\rangle, |j_1, -j_1 + 1\rangle, \ldots, |j_1, j_1 - 1\rangle, |j_1, j_1\rangle$ and $|j_2, -j_2\rangle, |j_2, -j_2 + 1\rangle, \ldots, |j_2, j_2 - 1\rangle, |j_2, j_2\rangle$.

The kets $|\, j_1, j_2;\, m_1, m_2\rangle$ form a complete and orthonormal basis. Using

$$\sum_{m_1 m_2} |\, j_1, j_2;\, m_1, m_2\rangle\langle j_1, j_2;\, m_1, m_2\,| = \left(\sum_{m_1} |\, j_1,\, m_1\rangle\langle j_1,\, m_1\,|\right)\left(\sum_{m_2} |\, j_2,\, m_2\rangle\langle j_2,\, m_2\,|\right), \tag{7.111}$$

and since $\{|\, j_1,\, m_1\rangle\}$ and $\{|\, j_2,\, m_2\rangle\}$ are complete (i.e., $\sum_{m_1} |\, j_1,\, m_1\rangle\langle j_1,\, m_1\,| = 1$) and orthonormal (i.e., $\langle j_1',\, m_1' \mid j_1,\, m_1\rangle = \delta_{j_1', j_1}\delta_{m_1', m_1}$ and similarly for $|\, j_2,\, m_2\rangle$), we see that the basis $\{|\, j_1, j_2;\, m_1, m_2\rangle\}$ is complete,

$$\sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |\, j_1, j_2;\, m_1, m_2\rangle\langle j_1, j_2;\, m_1, m_2\,| = 1, \tag{7.112}$$

and orthonormal,

$$\begin{aligned}\langle j_1', j_2';\, m_1', m_2' \mid j_1, j_2;\, m_1, m_2\rangle &= \langle j_1',\, m_1' \mid j_1,\, m_1\rangle\langle j_2',\, m_2' \mid j_2,\, m_2\rangle \\ &= \delta_{j_1', j_1}\delta_{j_2', j_2}\delta_{m_1', m_1}\delta_{m_2', m_2}.\end{aligned} \tag{7.113}$$

The basis $\{|\, j_1, j_2;\, m_1, m_2\rangle\}$ clearly spans the total space which is made of subspaces 1 and 2. From (7.106) we see that the dimension N of this space is equal to the product of the dimensions of the two subspaces spanned by $\{|\, j_1,\, m_1\rangle\}$ and $\{|\, j_2,\, m_2\rangle\}$:

$$N = (2j_1 + 1) \times (2j_2 + 1). \tag{7.114}$$

We can now introduce the *step* operators $\hat{J}_{1_\pm} = \hat{J}_{1_x} \pm i\hat{J}_{1_y}$ and $\hat{J}_{2_\pm} = \hat{J}_{2_x} \pm i\hat{J}_{2_y}$; their actions on $|j_1 j_2; m_1 m_2\rangle$ are given by

$$\hat{J}_{1\pm} |\, j_1, j_2;\, m_1, m_2\rangle = \hbar\sqrt{(j_1 \mp m_1)(j_1 \pm m_1 + 1)}\,|j_1, j_2;\, m_1 \pm 1, m_2\rangle, \tag{7.115}$$

$$\hat{J}_{2\pm} |\, j_1, j_2;\, m_1, m_2\rangle = \hbar\sqrt{(j_2 \mp m_2)(j_2 \pm m_2 + 1)}\,|j_1, j_2;\, m_1, m_2 \pm 1\rangle. \tag{7.116}$$

The problem of adding two angular momenta, $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$,

$$\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2, \tag{7.117}$$

consists of finding the eigenvalues and eigenvectors of $\hat{J}^2$ and $\hat{J}_z$ **in terms of the eigenvalues and eigenvectors of** $\hat{J}_1^2$, $\hat{J}_2^2$, $\hat{J}_{1_z}$, and $\hat{J}_{2_z}$. Since the matrices of $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ have in general different dimensions, the addition specified by (7.117) is not an addition of matrices; it is a symbolic addition.

By adding (7.99) and (7.100), we can easily ascertain that the components of $\hat{\vec{J}}$ satisfy the commutation relations of angular momentum:

$$\left[\hat{J}_x,\, \hat{J}_y\right] = i\hbar\hat{J}_z, \qquad \left[\hat{J}_y,\, \hat{J}_z\right] = i\hbar\hat{J}_x, \qquad \left[\hat{J}_z,\, \hat{J}_x\right] = i\hbar\hat{J}_y. \tag{7.118}$$

Note that $\hat{J}_1^2$, $\hat{J}_2^2$, $\hat{J}^2$, $\hat{J}_z$ jointly commute; this can be ascertained from the relation:

$$\hat{J}^2 = \hat{J}_1^2 + \hat{J}_2^2 + 2\hat{J}_{1_z}\hat{J}_{2_z} + \hat{J}_{1+}\hat{J}_{2-} + \hat{J}_{1-}\hat{J}_{2+}, \tag{7.119}$$

which leads to

$$\left[\hat{J}^2, \hat{J}_1^2\right] = \left[\hat{J}^2, \hat{J}_2^2\right] = 0, \tag{7.120}$$

and to

$$\left[\hat{J}^2, \hat{J}_z\right] = \left[\hat{J}_1^2, \hat{J}_z\right] = \left[\hat{J}_2^2, \hat{J}_z\right] = 0. \tag{7.121}$$

But in spite of the fact that $\left[\hat{J}^2, \hat{J}_z\right] = 0$, the operators $\hat{J}_{1_z}$ and $\hat{J}_{2_z}$ do not commute separately with $\hat{J}^2$:

$$\left[\hat{J}^2, \hat{J}_{1_z}\right] \neq 0, \qquad \left[\hat{J}^2, \hat{J}_{2_z}\right] \neq 0. \tag{7.122}$$

Now, since $\hat{J}_1^2, \hat{J}_2^2, \hat{J}^2, \hat{J}_z$ form a complete set of commuting operators, they can be diagonalized simultaneously by the same states; designating these joint eigenstates by $|j_1, j_2; j, m\rangle$, we have

$$\hat{J}_1^2 \mid j_1, j_2; j, m\rangle = j_1(j_1+1)\hbar^2 \mid j_1, j_2; j, m\rangle, \tag{7.123}$$

$$\hat{J}_2^2 \mid j_1, j_2; j, m\rangle = j_2(j_2+1)\hbar^2 \mid j_1, j_2; j, m\rangle, \tag{7.124}$$

$$\hat{J}^2 \mid j_1, j_2; j, m\rangle = j(j+1)\hbar^2 \mid j_1, j_2; j, m\rangle, \tag{7.125}$$

$$\hat{J}_z \mid j_1, j_2; j, m\rangle = m\hbar \mid j_1, j_2; j, m\rangle. \tag{7.126}$$

For every j, the number m has $(2j+1)$ allowed values: $m = -j, -j+1, \ldots, j-1, j$.

Since j_1 and j_2 are usually fixed, *we will be using, throughout the rest of this chapter, the shorthand notation* $\mid j, m\rangle$ *to abbreviate* $\mid j_1, j_2; j, m\rangle$. The set of vectors $\{\mid j, m\rangle\}$ form a complete and orthonormal basis:

$$\sum_j \sum_{m=-j}^{j} \mid j, m\rangle\langle j, m \mid = 1, \tag{7.127}$$

$$\langle j', m' \mid j, m\rangle = \delta_{j,j'}\delta_{m',m}. \tag{7.128}$$

The space where the total angular momentum $\hat{\vec{J}}$ operates is spanned by the basis $\{\mid j, m\rangle\}$; this space is known as a *product space.* It is important to know that this space is the same as the one spanned by $\{\mid j_1, j_2; m_1, m_2\rangle\}$; that is, the space which includes both subspaces 1 and 2. So the dimension of the space which is spanned by the basis $\{\mid j, m\rangle\}$ is also equal to $N = (2j_1+1) \times (2j_2+1)$ as specified by (7.114).

The issue now is to find the transformation that connects the bases $\{\mid j_1, j_2; m_1, m_2\rangle\}$ and $\{\mid j, m\rangle\}$.

7.3.1.1 Transformation between Bases: Clebsch–Gordan Coefficients

Let us now return to the addition of $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$. This problem consists in essence of obtaining the eigenvalues of $\hat{J}^2$ and $\hat{J}_z$ and of expressing the states $\mid j, m\rangle$ in terms of $\mid j_1, j_2; m_1, m_2\rangle$. We should mention that $\mid j, m\rangle$ is the state in which $\hat{J}^2$ and $\hat{J}_z$ have fixed values, $j(j+1)\hbar^2$ and $m\hbar$, but in general not a state in which the values of $\hat{J}_{1z}$ and $\hat{J}_{2z}$ are fixed; as for $\mid j_1, j_2; m_1, m_2\rangle$, it is the state in which $\hat{J}_1^2, \hat{J}_2^2, \hat{J}_{1z}$, and $\hat{J}_{2z}$ have fixed values.

The $\{|\, j_1, j_2;\, m_1, m_2\rangle\}$ and $\{|\, j,\, m\rangle\}$ bases can be connected by means of a transformation as follows. Inserting the identity operator as a sum over the complete basis $|\, j_1, j_2;\, m_1, m_2\rangle$, we can write

$$\begin{aligned} |\, j,\, m\rangle &= \left(\sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |\, j_1, j_2;\, m_1, m_2\rangle\langle j_1, j_2;\, m_1, m_2 \,| \right) |\, j,\, m\rangle \\ &= \sum_{m_1 m_2} \langle j_1, j_2;\, m_1, m_2 \,|\, j,\, m\rangle \,|\, j_1, j_2;\, m_1, m_2\rangle, \end{aligned} \tag{7.129}$$

where we have used the normalization condition (7.112); since the bases $\{|\, j_1, j_2;\, m_1, m_2\rangle\}$ and $\{|\, j,\, m\rangle\}$ are both *normalized*, this transformation must be *unitary*. The coefficients $\langle j_1, j_2;\, m_1, m_2 \,|\, j,\, m\rangle$, which depend only on the quantities j_1, j_2, j, m_1, m_2, and m, *are the matrix elements of the unitary transformation which connects the* $\{|\, j,\, m\rangle\}$ and $\{|\, j_1, j_2;\, m_1, m_2\rangle\}$ bases. These coefficients are called the *Clebsch–Gordan coefficients*.

The problem of angular momentum addition reduces then to finding the Clebsch–Gordan coefficients $\langle j_1, j_2;\, m_1, m_2 \,|\, j,\, m\rangle$. These coefficients are taken to be *real by convention*; hence

$$\boxed{\langle j_1, j_2;\, m_1, m_2 \,|\, j,\, m\rangle = \langle j,\, m \,|\, j_1, j_2;\, m_1, m_2\rangle.} \tag{7.130}$$

Using (7.112) and (7.128) we can infer the orthonormalization relation for the Clebsch–Gordan coefficients:

$$\sum_{m_1 m_2} \langle j',\, m' \,|\, j_1, j_2;\, m_1, m_2\rangle\langle j_1, j_2;\, m_1, m_2 \,|\, j,\, m\rangle = \delta_{j',j}\delta_{m',m}, \tag{7.131}$$

and since the Clebsch–Gordan coefficients are real, this relation can be rewritten as

$$\boxed{\sum_{m_1 m_2} \langle j_1, j_2;\, m_1, m_2 \,|\, j',\, m'\rangle\langle j_1, j_2;\, m_1, m_2 \,|\, j,\, m\rangle = \delta_{j',j}\delta_{m',m},} \tag{7.132}$$

which leads to

$$\boxed{\sum_{m_1 m_2} \langle j_1, j_2;\, m_1, m_2 \,|\, j,\, m\rangle^2 = 1.} \tag{7.133}$$

Likewise, we have

$$\boxed{\sum_{j} \sum_{m=-j}^{j} \langle j_1, j_2;\, m_1', m_2' \,|\, j,\, m\rangle\langle j_1, j_2;\, m_1, m_2 \,|\, j,\, m\rangle = \delta_{m_1', m_1}\delta_{m_2', m_2}} \tag{7.134}$$

and, in particular,

$$\boxed{\sum_{j}\sum_{m} \langle j_1, j_2;\, m_1, m_2 \,|\, j,\, m\rangle^2 = 1.} \tag{7.135}$$

7.3.1.2 Eigenvalues of $\hat{\vec{J}}^2$ and $\hat{J}_z$

Let us study how to find the eigenvalues of $\hat{\vec{J}}^2$ and $\hat{J}_z$ in terms of those of $\hat{\vec{J}}_1^2$, $\hat{\vec{J}}_2^2$, $\hat{J}_{1_z}$, and $\hat{J}_{2_z}$; that is, obtain j and m in terms of j_1, j_2, m_1 and m_2. First, since $\hat{J}_z = \hat{J}_{1_z} + \hat{J}_{2_z}$, we have $m = m_1 + m_2$.

Now, to find j in terms of j_1 and j_2, we proceed as follows. Since the maximum values of m_1 and m_2 are $m_{1max} = j_1$ and $m_{2max} = j_2$, we have $m_{max} = m_{1max} + m_{2max} = j_1 + j_2$; but since $|m| \leq j$, then $j_{max} = j_1 + j_2$.

Next, to find the minimum value j_{min} of j, we need to use the fact that there are a total of $(2j_1 + 1) \times (2j_2 + 1)$ eigenkets $|\, j,\ m\rangle$. To each value of j there correspond $(2j + 1)$ eigenstates $|\, j,\ m\rangle$, so we have

$$\sum_{j=j_{min}}^{j_{max}} (2j + 1) = (2j_1 + 1)(2j_2 + 1), \tag{7.136}$$

which leads to (see Example 7.2, page 454, for the proof)

$$j_{min}^2 = (j_1 - j_2)^2 \qquad \Longrightarrow \qquad j_{min} = |j_1 - j_2|. \tag{7.137}$$

Hence the allowed values of j are located within the range

$$\boxed{|j_1 - j_2| \leq j \leq j_1 + j_2.} \tag{7.138}$$

This expression can also be inferred from the well-known triangle relation[5]. So the allowed values of j proceed in integer steps according to

$$\boxed{j = |j_1 - j_2|,\ |j_1 - j_2| + 1,\ \ldots,\ j_1 + j_2 - 1,\ j_1 + j_2.} \tag{7.139}$$

Thus, for every j the allowed values of m are located within the range $-j \leq m \leq j$.

Note that the coefficient $\langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle$ vanishes unless $m_1 + m_2 = m$. This can be seen as follows: since $\hat{J}_z = \hat{J}_{1_z} + \hat{J}_{2_z}$, we have

$$\langle j_1, j_2;\ m_1, m_2 \mid \hat{J}_z - \hat{J}_{1_z} - \hat{J}_{2_z} \mid j,\ m\rangle = 0, \tag{7.140}$$

and since $\hat{J}_z \mid j,\ m\rangle = m\hbar \mid j,\ m\rangle$, $\langle j_1, j_2;\ m_1, m_2 \mid \hat{J}_{1_z} = m_1\hbar\langle j_1, j_2;\ m_1, m_2 \mid$, and $\langle j_1, j_2;\ m_1, m_2 \mid \hat{J}_{2_z} = m_2\hbar\langle j_1, j_2;\ m_1, m_2 \mid$, we can write

$$(m - m_1 - m_2)\langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle = 0, \tag{7.141}$$

which shows that $\langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle$ is not zero only when $m - m_1 - m_2 = 0$.

$$\boxed{\text{If} \quad m_1 + m_2 \neq m \quad \Longrightarrow \quad \langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle = 0.} \tag{7.142}$$

So, for the Clebsch–Gordan coefficient $\langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle$ not to be zero, we must simultaneously have

$$\boxed{m_1 + m_2 = m \qquad \text{and} \qquad |\, j_1 - j_2 \,| \leq j \leq j_1 + j_2.} \tag{7.143}$$

These are known as the *selection rules* for the Clebsch–Gordan coefficients.

---

[5]The length of the sum of two classical vectors, $\vec{A} + \vec{B}$, must be located between the sum and the difference of the lengths of the two vectors, $A + B$ and $|A - B|$, i.e., $|A - B| \leq |\vec{A} + \vec{B}| \leq A + B$.

Example 7.2
Starting from $\sum_{j=j_{min}}^{j_{max}}(2j+1) = (2j_1+1)(2j_2+1)$, prove (7.137).

Solution
Let us first work on the left-hand side of

$$\sum_{j=j_{min}}^{j_{max}}(2j+1) = (2j_1+1)(2j_2+1). \tag{7.144}$$

Since $j_{max} = j_1 + j_2$ we can write the left-hand side of this equation as an arithmetic sum which has $(j_{max} - j_{min} + 1) = [(j_1 + j_2 + 1) - j_{min}]$ terms:

$$\sum_{j=j_{min}}^{j_{max}}(2j+1) = (2j_{min}+1) + (2j_{min}+3) + (2j_{min}+5) + \cdots + [2(j_1+j_2)+1]. \tag{7.145}$$

To calculate this sum, we simply write it in the following two equivalent ways:

$$S = (2j_{min}+1) + (2j_{min}+3) + (2j_{min}+5) + \cdots + [2(j_1+j_2)+1], \tag{7.146}$$
$$S = [2(j_1+j_2)+1] + [2(j_1+j_2)-1] + [2(j_1+j_2)-3] + \cdots + (2j_{min}+1). \tag{7.147}$$

Adding these two series term by term, we obtain

$$2S = 2[(j_1+j_2+1)+j_{min}] + 2[(j_1+j_2+1)+j_{min}] + \cdots + 2[(j_1+j_2+1)+j_{min}]. \tag{7.148}$$

Since this expression has $(j_{max} - j_{min} + 1) = [(j_1 + j_2 + 1) - j_{min}]$ terms, we have

$$2S = 2[(j_1+j_2+1)+j_{min}][(j_1+j_2+1)-j_{min}]; \tag{7.149}$$

hence

$$S = [(j_1+j_2+1)+j_{min}][(j_1+j_2+1)-j_{min}] = (j_1+j_2+1)^2 - j_{min}^2. \tag{7.150}$$

Now, equating this expression with the right-hand side of (7.144), we obtain

$$(j_1+j_2+1)^2 - j_{min}^2 = (2j_1+1)(2j_2+1), \tag{7.151}$$

which in turn leads to

$$j_{min}^2 = (j_1-j_2)^2. \tag{7.152}$$

7.3.2 Calculation of the Clebsch–Gordan Coefficients

First, we should point out that the Clebsch–Gordan coefficients corresponding to the two limiting cases where $m_1 = j_1$, $m_2 = j_2$, $j = j_1 + j_2$, $m = j_1 + j_2$ and $m_1 = -j_1$, $m_2 = -j_2$, $j = j_1 + j_2$, $m = -(j_1 + j_2)$ are equal to one:

$$\boxed{\langle j_1, j_2;\ j_1, j_2 | (j_1+j_2),\ (j_1+j_2)\rangle = 1, \quad \langle j_1, j_2;\ -j_1, -j_2 | (j_1+j_2),\ -(j_1+j_2)\rangle = 1.} \tag{7.153}$$

These results can be inferred from (7.129), since $|(j_1 + j_2),\ (j_1 + j_2)\rangle$, and $|(j_1 + j_2),\ -(j_1 + j_2)\rangle$ have one element each:

$$|(j_1 + j_2),\ (j_1 + j_2)\rangle = \langle j_1, j_2;\ j_1,\ j_2|(j_1 + j_2),\ (j_1 + j_2)\rangle|j_1, j_2;\ j_1,\ j_2\rangle, \tag{7.154}$$

$$|(j_1 + j_2),\ -(j_1 + j_2)\rangle = \langle j_1, j_2;\ -j_1,\ -j_2|(j_1 + j_2),\ -(j_1 + j_2)\rangle|j_1, j_2;\ -j_1,\ -j_2\rangle, \tag{7.155}$$

where $|(j_1 + j_2),\ (j_1 + j_2)\rangle$, $|(j_1 + j_2),\ -(j_1 + j_2)\rangle$, $|j_1, j_2;\ j_1,\ j_2\rangle$, and $|j_1, j_2;\ -j_1,\ -j_2\rangle$ are all normalized.

The calculations of the other coefficients are generally more involved than the two limiting cases mentioned above. For this, we need to derive the recursion relations between the matrix elements of the unitary transformation between the $\{|\ j,\ m\rangle\}$ and $\{|j_1, j_2;\ m_1, m_2\rangle\}$ bases, since, when j_1, j_2 and j are fixed, the various Clebsch–Gordan coefficients are related to one another by means of recursion relations. To find the recursion relations, we need to evaluate the matrix elements $\langle j_1, j_2;\ m_1, m_2 \mid \hat{J}_\pm \mid j,\ m\rangle$ in two different ways. First, allow $\hat{J}_\pm$ to act to the right, i.e., on $|\ j,\ m\rangle$:

$$\langle j_1, j_2;\ m_1, m_2 \mid \hat{J}_\pm \mid j,\ m\rangle = \hbar\sqrt{(j \mp m)(j \pm m + 1)}\langle j_1, j_2;\ m_1, m_2 \mid j,\ m \pm 1\rangle. \tag{7.156}$$

Second, make $\hat{J}_\pm = \hat{J}_{1\pm} + \hat{J}_{2\pm}$ act to the left[6], i.e., on $\langle j_1, j_2;\ m_1, m_2\ |$:

$$\begin{aligned}\langle j_1, j_2;\ m_1, m_2 \mid \hat{J}_\pm \mid j,\ m\rangle &= \hbar\sqrt{(j_1 \pm m_1)(j_1 \mp m_1 + 1)}\langle j_1, j_2;\ m_1 \mp 1, m_2 \mid j,\ m\rangle \\ &\quad + \hbar\sqrt{(j_2 \pm m_2)(j_2 \mp m_2 + 1)}\langle j_1, j_2;\ m_1, m_2 \mp 1 \mid j,\ m\rangle.\end{aligned} \tag{7.157}$$

Equating (7.156) and (7.157) we obtain the desired recursion relations for the Clebsch–Gordan coefficients:

$$\boxed{\begin{aligned}&\sqrt{(j \mp m)(j \pm m + 1)}\ \langle j_1, j_2;\ m_1, m_2 \mid j,\ m \pm 1\rangle \\ &\qquad = \sqrt{(j_1 \pm m_1)(j_1 \mp m_1 + 1)}\langle j_1, j_2; m_1 \mp 1, m_2 \mid j,\ m\rangle \\ &\qquad\quad + \sqrt{(j_2 \pm m_2)(j_2 \mp m_2 + 1)}\langle j_1, j_2;\ m_1, m_2 \mp 1 \mid j,\ m\rangle.\end{aligned}} \tag{7.158}$$

These relations, together with the orthonormalization relation (7.133), determine all Clebsch–Gordan coefficients for any given values of j_1, j_2, and j. To see this, let us substitute $m_1 = j_1$ and $m = j$ into the lower part of (7.158). Since m_2 can be equal only to $m_2 = j - j_1 - 1$, we obtain

$$\begin{aligned}\sqrt{2j}\langle j_1, j_2;\ j_1, (j - j_1 - 1)|j,\ j - 1\rangle\ &= \sqrt{(j_2 - j + j_1 + 1)(j_2 + j - j_1)} \\ &\quad \times\langle j_1, j_2;\ j_1, (j - j_1)|j,\ j\rangle.\end{aligned} \tag{7.159}$$

Thus, knowing $\langle j_1, j_2;\ j_1, (j - j_1)|j,\ j\rangle$, we can determine $\langle j_1, j_2;\ j_1, (j - j_1 - 1)|j,\ j - 1\rangle$. In addition, substituting $m_1 = j_1$, $m = j - 1$ and $m_2 = j - j_1$ into the upper part of (7.158), we end up with

$$\begin{aligned}\sqrt{2j}\,\langle j_1, j_2;\ j_1, (j - j_1) \mid j,\ j\rangle &= \sqrt{2j_1}\langle j_1, j_2;\ (j_1 - 1), (j - j_1) \mid j,\ j - 1\rangle \\ &\quad + \sqrt{(j_2 + j - j_1)(j_2 - j + j_1 + 1)}\langle j_1, j_2;\ j_1, (j - j_1 - 1) \mid j,\ j - 1\rangle.\end{aligned} \tag{7.160}$$

[6]Recall that $\langle j_1, j_2;\ m_1, m_2|\hat{J}_{1\pm} = \hbar\sqrt{(j_1 \pm m_1)(j_1 \mp m_1 + 1)}\langle j_1, j_2;\ m_1 \mp 1, m_2|$.

Thus knowing $\langle j_1, j_2;\ j_1, (j-j_1) \mid j,\ j\rangle$ and $\langle j_1, j_2;\ j_1, (j-j_1-1) \mid j,\ j-1\rangle$, we can determine $\langle j_1, j_2;\ (j_1-1), (j-j_1) \mid j,\ j-1\rangle$. Repeated application of the recursion relation (7.158) will determine all the other Clebsch–Gordan coefficients, provided we know only one of them: $\langle j_1, j_2;\ j_1, (j-j_1) \mid j,\ j\rangle$. As for the absolute value of this coefficient, it can be determined from the normalization condition (7.132). Thus, the recursion relation (7.158), in conjunction with the normalization condition (7.132), determines all the Clebsch–Gordan coefficients except for a sign. But how does one determine this sign?

The convention, known as the phase convention, is to consider $\langle j_1, j_2;\ j_1, (j-j_1) | j,\ j\rangle$ to be *real* and *positive*. This phase convention implies that

$$\boxed{\langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle = (-1)^{j-j_1-j_2}\langle j_2, j_1;\ m_2, m_1 \mid j,\ m\rangle;} \tag{7.161}$$

hence

$$\boxed{\begin{aligned}\langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle &= (-1)^{j-j_1-j_2}\langle j_1, j_2;\ -m_1,\ -m_2 | j,\ -m\rangle \\ &= \langle j_2, j_1;\ -m_2,\ -m_1 | j,\ -m\rangle.\end{aligned}} \tag{7.162}$$

Note that, since all the Clebsch–Gordan coefficients are obtained from a single coefficient $\langle j_1, j_2;\ j_1, (j-j_1) | j,\ j\rangle$, and since this coefficient is real, all other Clebsch–Gordan coefficients must also be real numbers.

Following the same method that led to (7.158) from $\langle j_1, j_2;\ m_1, m_2 \mid \hat{J}_\pm \mid j,\ m\rangle$, we can show that a calculation of $\langle j_1, j_2;\ m_1, m_2 \mid \hat{J}_\pm | j\ m, \mp 1\rangle$ leads to the following recursion relation:

$$\boxed{\begin{aligned}&\sqrt{(j \mp m + 1)(j \pm m)}\ \langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle \\ &\qquad = \sqrt{(j_1 \pm m_1)(j_1 \mp m_1 + 1)}\langle j_1, j_2;\ m_1 \mp 1,\ m_2 | j,\ m \mp 1\rangle \\ &\qquad\quad + \sqrt{(j_2 \pm m_2)(j_2 \mp m_2 + 1)}\langle j_1, j_2;\ m_1, m_2 \mp 1 | j,\ m \mp 1\rangle.\end{aligned}} \tag{7.163}$$

We can use the recursion relations (7.158) and (7.163) to obtain the values of the various Clebsch–Gordan coefficients. For instance, if we insert $m_1 = j_1$, $m_2 = j_2 - 1$, $j = j_1 + j_2$, and $m = j_1 + j_2$ into the lower sign of (7.158), we obtain

$$\langle j_1, j_2;\ j_1, (j_2-1) | (j_1+j_2),\ (j_1+j_2-1)\rangle = \sqrt{\frac{j_2}{j_1+j_2}}. \tag{7.164}$$

Similarly, a substitution of $m_1 = j_1 - 1$, $m_2 = j_2$, $j = j_1 + j_2$, and $m = j_1 + j_2$ into the lower sign of (7.158) leads to

$$\langle j_1, j_2;\ (j_1-1), j_2 | (j_1+j_2),\ (j_1+j_2-1)\rangle = \sqrt{\frac{j_1}{j_1+j_2}}. \tag{7.165}$$

We can also show that

$$\langle j, 1;\ m, 0 \mid j,\ m\rangle = \frac{m}{\sqrt{j(j+1)}}, \qquad \langle j, 0;\ m, 0 \mid j,\ m\rangle = 1. \tag{7.166}$$

Example 7.3

(a) Find the Clebsch–Gordan coefficients associated with the coupling of the spins of the electron and the proton of a hydrogen atom in its ground state.

(b) Find the transformation matrix which is formed by the Clebsch–Gordan coefficients. Verify that this matrix is unitary.

Solution

In their ground states the proton and electron have no orbital angular momenta. Thus, the total angular momentum of the atom is obtained by simply adding the spins of the proton and electron.

This is a simple example to illustrate the general formalism outlined in this section. Since $j_1 = \frac{1}{2}$ and $j_2 = \frac{1}{2}$, j has two possible values $j = 0,\ 1$. When $j = 0$, there is only a single state $|\, j,\ m\rangle = |\, 0,\ 0\rangle$; this is called the *spin singlet*. On the other hand, there are three possible values of $m = -1,\ 0,\ 1$ for the case $j = 1$; this corresponds to a *spin triplet state* $|\, 1,\ -1\rangle, |\, 1,\ 0\rangle, |\, 1,\ 1\rangle$.

From (7.129), we can express the states $|\, j,\ m\rangle$ in terms of $|\, \frac{1}{2}, \frac{1}{2};\ m_1, m_2\rangle$ as follows:

$$|\, j,\ m\rangle = \sum_{m_1=-1/2}^{1/2} \sum_{m_2=-1/2}^{1/2} \langle \tfrac{1}{2}, \tfrac{1}{2};\ m_1, m_2 \,|\, j,\ m\rangle \,|\, \tfrac{1}{2}, \tfrac{1}{2};\ m_1, m_2\rangle, \tag{7.167}$$

which, when applied to the two cases $j = 0$ and $j = 1$, leads to

$$|\, 0,\ 0\rangle = \langle \tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, -\tfrac{1}{2} \,|\, 0,\ 0\rangle \left|\tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, -\tfrac{1}{2}\right\rangle + \langle \tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, \tfrac{1}{2} \,|\, 0,\ 0\rangle \left|\tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, \tfrac{1}{2}\right\rangle, \tag{7.168}$$

$$|\, 1,\ 1\rangle = \langle \tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, \tfrac{1}{2} \,|\, 1,\ 1\rangle \left|\tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, \tfrac{1}{2}\right\rangle, \tag{7.169}$$

$$|\, 1,\ 0\rangle = \langle \tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, -\tfrac{1}{2} \,|\, 1,\ 0\rangle \left|\tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, -\tfrac{1}{2}\right\rangle + \langle \tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, \tfrac{1}{2} \,|\, 1,\ 0\rangle \left|\tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, \tfrac{1}{2}\right\rangle, \tag{7.170}$$

$$|\, 1,\ -1\rangle = \langle \tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, -\tfrac{1}{2} \,|\, 1,\ -1\rangle \left|\tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle. \tag{7.171}$$

To calculate the Clebsch–Gordan coefficients involved in (7.168)–(7.171), we are going to adopt two separate approaches: the first approach uses the recursion relations (7.158) and (7.163), while the second uses the algebra of angular momentum.

First approach: using the recursion relations

First, to calculate the two coefficients $\langle \frac{1}{2}, \frac{1}{2};\ \pm\frac{1}{2}, \mp\frac{1}{2} \,|\, 0,\ 0\rangle$ involved in (7.168), we need, on the one hand, to substitute $j = 0, m = 0, m_1 = m_2 = \frac{1}{2}$ into the upper sign relation of (7.158):

$$\langle \tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, \tfrac{1}{2} \,|\, 0,\ 0\rangle = -\langle \tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, -\tfrac{1}{2} \,|\, 0,\ 0\rangle. \tag{7.172}$$

On the other hand, the substitution of $j = 0$ and $m = 0$ into (7.133) yields

$$\langle \tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, \tfrac{1}{2} \,|\, 0,\ 0\rangle^2 + \langle \tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, -\tfrac{1}{2} \,|\, 0,\ 0\rangle^2 = 1 \tag{7.173}$$

Combining (7.172) and (7.173) we end up with

$$\langle \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, -\tfrac{1}{2} \mid 0, 0 \rangle = \pm \frac{1}{\sqrt{2}}. \tag{7.174}$$

The sign of $\langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} | 0, 0 \rangle$ has to be positive because, according to the phase convention, the coefficient $\langle j_1, j_2; j_1, (j - j_1) \mid j, j \rangle$ is positive; hence

$$\langle \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, -\tfrac{1}{2} \mid 0, 0 \rangle = \frac{1}{\sqrt{2}}. \tag{7.175}$$

As for $\langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \mid 0, 0 \rangle$, its value can be inferred from (7.172) and (7.175):

$$\langle \tfrac{1}{2}, \tfrac{1}{2}; -\tfrac{1}{2}, \tfrac{1}{2} \mid 0, 0 \rangle = -\frac{1}{\sqrt{2}}. \tag{7.176}$$

Second, the calculation of the coefficients involved in (7.169) to (7.171) goes as follows. The orthonormalization relation (7.133) leads to

$$\langle \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, \tfrac{1}{2} \mid 1, 1 \rangle^2 = 1, \qquad \langle \tfrac{1}{2}, \tfrac{1}{2}; -\tfrac{1}{2}, -\tfrac{1}{2} \mid 1, -1 \rangle^2 = 1, \tag{7.177}$$

and since $\langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \mid 1, 1 \rangle$ and $\langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} \mid 1, -1 \rangle$ are both real and positive, we have

$$\langle \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, \tfrac{1}{2} \mid 1, 1 \rangle = 1, \qquad \langle \tfrac{1}{2}, \tfrac{1}{2}; -\tfrac{1}{2}, -\tfrac{1}{2} \mid 1, -1 \rangle = 1. \tag{7.178}$$

As for the coefficients $\langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \mid 1, 0 \rangle$ and $\langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \mid 1, 0 \rangle$, they can be extracted by setting $j = 1, m = 0, m_1 = \frac{1}{2}, m_2 = -\frac{1}{2}$ and $j = 1, m = 0, m_1 = -\frac{1}{2}, m_2 = \frac{1}{2}$, respectively, into the lower sign case of (7.163):

$$\sqrt{2} \langle \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, -\tfrac{1}{2} \mid 1, 0 \rangle = \langle \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, \tfrac{1}{2} \mid 1, 1 \rangle. \tag{7.179}$$

$$\sqrt{2} \langle \tfrac{1}{2}, \tfrac{1}{2}; -\tfrac{1}{2}, \tfrac{1}{2} \mid 1, 0 \rangle = \langle \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, \tfrac{1}{2} \mid 1, 1 \rangle, \tag{7.180}$$

Combining (7.178) with (7.180) and (7.179), we find

$$\langle \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, -\tfrac{1}{2} \mid 1, 0 \rangle = \langle \tfrac{1}{2}, \tfrac{1}{2}; -\tfrac{1}{2}, \tfrac{1}{2} \mid 1, 0 \rangle = \frac{1}{\sqrt{2}}. \tag{7.181}$$

Finally, substituting the Clebsch–Gordan coefficients (7.175) and (7.176) into (7.168), and substituting (7.178) and (7.181) into (7.169) through (7.171), we end up with

$$| 0, 0 \rangle = -\frac{1}{\sqrt{2}} \left| \tfrac{1}{2}, \tfrac{1}{2}; -\tfrac{1}{2}, \tfrac{1}{2} \right\rangle + \frac{1}{\sqrt{2}} \left| \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, -\tfrac{1}{2} \right\rangle, \tag{7.182}$$

$$| 1, 1 \rangle = \left| \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, \tfrac{1}{2} \right\rangle, \tag{7.183}$$

$$| 1, 0 \rangle = \frac{1}{\sqrt{2}} \left| \tfrac{1}{2}, \tfrac{1}{2}; -\tfrac{1}{2}, \tfrac{1}{2} \right\rangle + \frac{1}{\sqrt{2}} \left| \tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, -\tfrac{1}{2} \right\rangle, \tag{7.184}$$

$$| 1, -1 \rangle = \left| \tfrac{1}{2}, \tfrac{1}{2}; -\tfrac{1}{2}, -\tfrac{1}{2} \right\rangle. \tag{7.185}$$

Note that the singlet state $|0,\ 0\rangle$ is antisymmetric, whereas the triplet states $|1,\ -1\rangle$, $|1,\ 0\rangle$, and $|1,\ 1\rangle$ are symmetric.

Second approach: using angular momentum algebra
Beginning with $j = 1$, and since $|\ 1,\ 1\rangle$ and $|\ \frac{1}{2}, \frac{1}{2};\ \frac{1}{2}, \frac{1}{2}\rangle$ are both normalized, equation (7.169) leads to

$$\langle \tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, \tfrac{1}{2}\ |\ 1,\ 1\rangle^2 = 1. \tag{7.186}$$

From the phase convention, which states that $\langle j_1, j_2;\ j_1, (j - j_1)|j,\ j\rangle$ must be positive, we see that $\langle \frac{1}{2}, \frac{1}{2};\ \frac{1}{2}, \frac{1}{2}\ |\ 1,\ 1\rangle = 1$, and hence

$$|\ 1,\ 1\rangle = \left|\tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, \tfrac{1}{2}\right\rangle. \tag{7.187}$$

Now, to find the Clebsch–Gordan coefficients in $|\ 1,\ 0\rangle$, we simply apply $\hat{J}_-$ on $|\ 1,\ 1\rangle$:

$$\hat{J}_-\ |\ 1,\ 1\rangle = (\hat{J}_{1-} + \hat{J}_{2-})\left|\tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, \tfrac{1}{2}\right\rangle, \tag{7.188}$$

which leads to

$$|\ 1,\ 0\rangle = \frac{1}{\sqrt{2}}\left|\tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, \tfrac{1}{2}\right\rangle + \frac{1}{\sqrt{2}}\left|\tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, -\tfrac{1}{2}\right\rangle, \tag{7.189}$$

hence $\langle \frac{1}{2}, \frac{1}{2};\ -\frac{1}{2}, \frac{1}{2}\ |\ 1,\ 0\rangle = 1/\sqrt{2}$ and $\langle \frac{1}{2}, \frac{1}{2};\ \frac{1}{2}, -\frac{1}{2}\ |\ 1,\ 0\rangle = 1/\sqrt{2}$. Next, applying $\hat{J}_-$ on (7.189), we get

$$|\ 1,\ -1\rangle = \left|\tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle. \tag{7.190}$$

Finally, to find $|\ 0,\ 0\rangle$, we proceed in two steps: first, since

$$|\ 0,\ 0\rangle = a\left|\tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, \tfrac{1}{2}\right\rangle + b\left|\tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, -\tfrac{1}{2}\right\rangle, \tag{7.191}$$

where $a = \langle \frac{1}{2}, \frac{1}{2};\ -\frac{1}{2}, \frac{1}{2}\ |\ 0,\ 0\rangle$ and $b = \langle \frac{1}{2}, \frac{1}{2};\ \frac{1}{2}, -\frac{1}{2}\ |\ 0,\ 0\rangle$, a combination of (7.189) with (7.191) leads to

$$\langle 0,\ 0\ |\ 1,\ 0\rangle = \frac{a}{\sqrt{2}} + \frac{b}{\sqrt{2}} = 0; \tag{7.192}$$

second, since $|\ 0,\ 0\rangle$ is normalized, we have

$$\langle 0,\ 0\ |\ 0,\ 0\rangle = a^2 + b^2 = 1. \tag{7.193}$$

Combining (7.192) and (7.193), and since $\langle \frac{1}{2}, \frac{1}{2};\ \frac{1}{2}, -\frac{1}{2}\ |\ 0,\ 0\rangle$ must be positive, we obtain $a = \langle \frac{1}{2}, \frac{1}{2};\ -\frac{1}{2}, \frac{1}{2}\ |\ 0,\ 0\rangle = -1/\sqrt{2}$ and $b = \langle \frac{1}{2}, \frac{1}{2};\ \frac{1}{2}, -\frac{1}{2}\ |\ 0,\ 0\rangle = 1/\sqrt{2}$. Inserting these values into (7.191) we obtain

$$|\ 0,\ 0\rangle = -\frac{1}{\sqrt{2}}\left|\tfrac{1}{2}, \tfrac{1}{2};\ -\tfrac{1}{2}, \tfrac{1}{2}\right\rangle + \frac{1}{\sqrt{2}}\left|\tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{2}, -\tfrac{1}{2}\right\rangle. \tag{7.194}$$

(b) Writing (7.182) to (7.185) in a matrix form:

$$\begin{pmatrix} |0,\ 0\rangle \\ |1,\ 1\rangle \\ |1,\ 0\rangle \\ |1,-1\rangle \end{pmatrix} = \begin{pmatrix} 0 & 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} |\ \frac{1}{2}, \frac{1}{2};\ \frac{1}{2},\ \frac{1}{2}\rangle \\ |\ \frac{1}{2}, \frac{1}{2};\ \frac{1}{2}, -\frac{1}{2}\rangle \\ |\ \frac{1}{2}, \frac{1}{2};\ -\frac{1}{2},\ \frac{1}{2}\rangle \\ |\ \frac{1}{2}, \frac{1}{2};\ -\frac{1}{2}, -\frac{1}{2}\rangle \end{pmatrix}, \tag{7.195}$$

we see that the elements of the transformation matrix

$$U = \begin{pmatrix} 0 & 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{7.196}$$

which connects the $\{|\ j,\ m\rangle\}$ vectors to their $\{|\ j_1, j_2;\ m_1, m_2\rangle\}$ counterparts, are given by the Clebsch–Gordan coefficients derived above. Inverting (7.195) we obtain

$$\begin{pmatrix} |\frac{1}{2},\frac{1}{2};\ \frac{1}{2},\ \frac{1}{2}\rangle \\ |\frac{1}{2},\frac{1}{2};\ \frac{1}{2},-\frac{1}{2}\rangle \\ |\frac{1}{2},\frac{1}{2};-\frac{1}{2},\ \frac{1}{2}\rangle \\ |\frac{1}{2},\frac{1}{2};-\frac{1}{2},-\frac{1}{2}\rangle \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} |0,\ 0\rangle \\ |1,\ 1\rangle \\ |1,\ 0\rangle \\ |1,-1\rangle \end{pmatrix}. \tag{7.197}$$

From (7.195) and (7.197) we see that the transformation matrix U is unitary since $U^{-1} = U^{\dagger}$.

7.3.3 Coupling of Orbital and Spin Angular Momenta

We consider here an important application of the formalism of angular momenta addition to the coupling of an orbital and a spin angular momentum: $\vec{J} = \vec{L} + \vec{S}$. In particular, we want to find Clebsch–Gordan coefficients associated with this coupling for a spin $s = \frac{1}{2}$ particle. In this case we have: $j_1 = l$ (integer), $m_1 = m_l$, $j_2 = s = \frac{1}{2}$, and $m_2 = m_s = \pm\frac{1}{2}$. The allowed values of j as given by (7.138) are located within the interval $|l - \frac{1}{2}| \leq j \leq |l + \frac{1}{2}|$. If $l = 0$ the problem would be obvious: the particle would have only spin and no orbital angular momentum. But if $l > 0$ then j can take only two possible values $j = l \pm \frac{1}{2}$. There are $2(l+1)$ states $\{|\ l + \frac{1}{2}, m\rangle\}$ corresponding to the case $j = l + 1/2$ and $2l$ states $\{|\ l - \frac{1}{2}, m\rangle\}$ corresponding to $j = l - \frac{1}{2}$. Let us study in detail each one of these two cases.

Case $j = l + 1/2$
Applying the relation (7.129) to the case where $j = l + \frac{1}{2}$, we have

$$\begin{aligned} \left|l+\frac{1}{2},m\right\rangle &= \sum_{m_l=-l}^{l}\sum_{m_2=-1/2}^{1/2} \left\langle l,\frac{1}{2};\ m_l,m_2\middle|l+\frac{1}{2},\ m\right\rangle \left|l,\frac{1}{2};\ m_l,m_2\right\rangle \\ &= \sum_{m_l} \left\langle l,\frac{1}{2};\ m_l,-\frac{1}{2}\middle|l+\frac{1}{2},\ m\right\rangle \left|l,\frac{1}{2};\ m_l,-\frac{1}{2}\right\rangle \\ &\quad + \sum_{m_l} \left\langle l,\frac{1}{2};\ m_l,\frac{1}{2}\middle|l+\frac{1}{2},\ m\right\rangle \left|l,\frac{1}{2};\ m_l,\frac{1}{2}\right\rangle. \end{aligned} \tag{7.198}$$

Using the selection rule $m_l + m_2 = m$ or $m_l = m - m_2$, we can rewrite (7.198) as follows:

$$\begin{aligned} \left|l+\frac{1}{2},m\right\rangle &= \left\langle l,\frac{1}{2};\ m+\frac{1}{2},-\frac{1}{2}\middle|l+\frac{1}{2},m\right\rangle \left|l,\frac{1}{2};\ m+\frac{1}{2},-\frac{1}{2}\right\rangle \\ &\quad + \left\langle l,\frac{1}{2};\ m-\frac{1}{2},\frac{1}{2}\middle|l+\frac{1}{2},m\right\rangle \left|l,\frac{1}{2};\ m-\frac{1}{2},\frac{1}{2}\right\rangle. \end{aligned} \tag{7.199}$$

We need now to calculate $\langle l, \frac{1}{2};\ m+\frac{1}{2}, -\frac{1}{2}|l+\frac{1}{2},\ m\rangle$ and $\langle l, \frac{1}{2};\ m-\frac{1}{2}, \frac{1}{2}|l+\frac{1}{2},\ m\rangle$. We begin with the calculation of $\langle l, \frac{1}{2};\ m+\frac{1}{2}, -\frac{1}{2}|l+\frac{1}{2},\ m\rangle$. Substituting $j = l+\frac{1}{2}$, $j_1 = l$, $j_2 = \frac{1}{2}$, $m_1 = m+\frac{1}{2}$, $m_2 = -\frac{1}{2}$ into the upper sign case of (7.163), we obtain

$$\sqrt{\left(l-m+\frac{3}{2}\right)\left(l+m+\frac{1}{2}\right)}\left\langle l,\frac{1}{2};\ m+\frac{1}{2},-\frac{1}{2}\middle|l+\frac{1}{2},m\right\rangle = \sqrt{\left(l+m+\frac{1}{2}\right)\left(l-m+\frac{1}{2}\right)}\left\langle l,\frac{1}{2};\ m-\frac{1}{2},-\frac{1}{2}\middle|l+\frac{1}{2},m-1\right\rangle \tag{7.200}$$

or

$$\left\langle l,\frac{1}{2};\ m+\frac{1}{2},-\frac{1}{2}\middle|l+\frac{1}{2},\ m\right\rangle = \sqrt{\frac{l-m+1/2}{l-m+3/2}}\left\langle l,\frac{1}{2};\ m-\frac{1}{2},-\frac{1}{2}\middle|l+\frac{1}{2},\ m-1\right\rangle. \tag{7.201}$$

By analogy with $\langle l, \frac{1}{2};\ m+\frac{1}{2}, -\frac{1}{2}|l+\frac{1}{2},\ m\rangle$ we can express the Clebsch–Gordan coefficient $\langle l, \frac{1}{2};\ m-\frac{1}{2}, -\frac{1}{2}|l+\frac{1}{2},\ m-1\rangle$ in terms of $\langle l, \frac{1}{2};\ m-\frac{3}{2}, -\frac{1}{2}|l+\frac{1}{2},\ m-2\rangle$ as follows:

$$\left\langle l,\frac{1}{2};\ m-\frac{1}{2},-\frac{1}{2}\middle|l+\frac{1}{2},\ m-1\right\rangle = \sqrt{\frac{l-m+3/2}{l-m+5/2}}\left\langle l,\frac{1}{2};\ m-\frac{3}{2},-\frac{1}{2}\middle|l+\frac{1}{2},\ m-2\right\rangle, \tag{7.202}$$

which, when inserted into (7.201), leads to

$$\left\langle l,\frac{1}{2};\ m+\frac{1}{2},-\frac{1}{2}\middle|l+\frac{1}{2},\ m\right\rangle = \sqrt{\frac{l-m+1/2}{l-m+3/2}}\sqrt{\frac{l-m+3/2}{l-m+5/2}} \times\left\langle l,\frac{1}{2};\ m-\frac{3}{2},-\frac{1}{2}\middle|l+\frac{1}{2},m-2\right\rangle. \tag{7.203}$$

We can continue this procedure until m reaches its lowest values, $-l-\frac{1}{2}$:

$$\left\langle l,\frac{1}{2};\ m+\frac{1}{2},-\frac{1}{2}\middle|l+\frac{1}{2},m\right\rangle = \sqrt{\frac{l-m+1/2}{l-m+3/2}}\sqrt{\frac{l-m+3/2}{l-m+5/2}} \times\cdots\times\sqrt{\frac{2l}{2l+1}}\left\langle l,\frac{1}{2};\ -l,-\frac{1}{2}\middle|l+\frac{1}{2},-l-\frac{1}{2}\right\rangle, \tag{7.204}$$

or

$$\left\langle l,\frac{1}{2};\ m+\frac{1}{2},-\frac{1}{2}\middle|l+\frac{1}{2},\ m\right\rangle = \sqrt{\frac{l-m+1/2}{2l+1}}\left\langle l,\frac{1}{2};\ -l,-\frac{1}{2}\middle|l+\frac{1}{2},\ -l-\frac{1}{2}\right\rangle. \tag{7.205}$$

From (7.133) we can easily obtain $\left\langle l,\frac{1}{2};\ -l,-\frac{1}{2}\middle|l+\frac{1}{2},-l-\frac{1}{2}\right\rangle^2 = 1$, and since this coefficient is real we have $\left\langle l,\frac{1}{2};\ -l,-\frac{1}{2}\middle|l+\frac{1}{2},-l-\frac{1}{2}\right\rangle = 1$. Inserting this value into (7.205) we end up with

$$\boxed{\left\langle l,\frac{1}{2};\ m+\frac{1}{2},-\frac{1}{2}\middle|l+\frac{1}{2},\ m\right\rangle = \sqrt{\frac{l-m+1/2}{2l+1}}.} \tag{7.206}$$

Now we turn to the calculation of the second coefficient, $\langle l, \frac{1}{2};\ m-\frac{1}{2}, \frac{1}{2}|l+\frac{1}{2},\ m\rangle$, involved in (7.199). We can perform this calculation in two different ways. The first method consists of following the same procedure adopted above to find $\langle l, \frac{1}{2};\ m+\frac{1}{2}, -\frac{1}{2}|l+\frac{1}{2}, m\rangle$. For this, we need only to substitute $j = l+\frac{1}{2}, j_1 = l, j_2 = \frac{1}{2}, m_1 = m-\frac{1}{2}, m_2 = \frac{1}{2}$ in the lower sign case of (7.163) and work our way through. A second, simpler method consists of substituting (7.206) into (7.199) and then calculating the norm of the resulting equation:

$$1 = \frac{l-m+1/2}{2l+1} + \left\langle l, \frac{1}{2};\ m-\frac{1}{2}, \frac{1}{2}\middle| l+\frac{1}{2},\ m\right\rangle^2, \tag{7.207}$$

where we have used the facts that the three kets $|l+\frac{1}{2}, m\rangle$ and $\left|l, \frac{1}{2};\ m\pm\frac{1}{2}, \mp\frac{1}{2}\right\rangle$ are normalized. Again, since $\langle l, \frac{1}{2};\ m-\frac{1}{2}, \frac{1}{2}|l+\frac{1}{2}, m\rangle$ is real, (7.207) leads to

$$\boxed{\left\langle l, \frac{1}{2};\ m-\frac{1}{2}, \frac{1}{2}\middle| l+\frac{1}{2},\ m\right\rangle = \sqrt{\frac{l+m+1/2}{2l+1}}.} \tag{7.208}$$

A combination of (7.199), (7.206), and (7.208) yields

$$\left|l+\frac{1}{2},\ m\right\rangle = \sqrt{\frac{l-m+1/2}{2l+1}}\left|l, \frac{1}{2};\ m+\frac{1}{2}, -\frac{1}{2}\right\rangle + \sqrt{\frac{l+m+1/2}{2l+1}}\left|l, \frac{1}{2};\ m-\frac{1}{2}, \frac{1}{2}\right\rangle, \tag{7.209}$$

where the possible values of m are given by

$$m = -l-\frac{1}{2}, -l+\frac{1}{2}, -l+\frac{3}{2}, \ldots, l-\frac{3}{2}, l-\frac{1}{2}, l+\frac{1}{2}. \tag{7.210}$$

Case $j = l - 1/2$

There are $2l$ states, $\{|l-\frac{1}{2},\ m\rangle\}$, corresponding to $j = l-\frac{1}{2}$; these are $\left|l-\frac{1}{2},\ -l+\frac{1}{2}\right\rangle, \left|l-\frac{1}{2},\ -l+\frac{3}{2}\right\rangle, \ldots, \left|l-\frac{1}{2},\ l-\frac{1}{2}\right\rangle$. Using (7.129) we write any state $\left|l-\frac{1}{2}, m\right\rangle$ as

$$\begin{aligned}\left|l-\frac{1}{2}, m\right\rangle = &\left\langle l, \frac{1}{2};\ m+\frac{1}{2}, -\frac{1}{2}\middle| l-\frac{1}{2},\ m\right\rangle\left|l, \frac{1}{2};\ m+\frac{1}{2}, -\frac{1}{2}\right\rangle \\ &+ \left\langle l, \frac{1}{2};\ m-\frac{1}{2}, \frac{1}{2}\middle| l-\frac{1}{2},\ m\right\rangle\left|l, \frac{1}{2};\ m-\frac{1}{2}, \frac{1}{2}\right\rangle.\end{aligned} \tag{7.211}$$

The two Clebsch–Gordan coefficients involved in this equation can be calculated by following the same method that we adopted above for the case $j = l+\frac{1}{2}$. Thus, we can ascertain that $|l-\frac{1}{2}, m\rangle$ is given by

$$\left|l-\frac{1}{2},\ m\right\rangle = \sqrt{\frac{l+m+1/2}{2l+1}}\left|l, \frac{1}{2};\ m+\frac{1}{2}, -\frac{1}{2}\right\rangle - \sqrt{\frac{l-m+1/2}{2l+1}}\left|l, \frac{1}{2};\ m-\frac{1}{2}, \frac{1}{2}\right\rangle, \tag{7.212}$$

where

$$m = -l+\frac{1}{2}, -l+\frac{3}{2}, \ldots, l-\frac{3}{2}, l-\frac{1}{2}. \tag{7.213}$$

We can combine (7.209) and (7.212) into

$$\boxed{\left|l \pm \frac{1}{2},\, m\right\rangle = \sqrt{\frac{l \mp m + \frac{1}{2}}{2l+1}}\left|l, \frac{1}{2};\, m + \frac{1}{2}, -\frac{1}{2}\right\rangle \pm \sqrt{\frac{l \pm m + \frac{1}{2}}{2l+1}}\left|l, \frac{1}{2};\, m - \frac{1}{2}, \frac{1}{2}\right\rangle.} \tag{7.214}$$

Illustration on a particle with $l = 1$

As an illustration of the formalism worked out above, we consider the particular case of $l = 1$. Inserting $l = 1$ and $m = \frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}$ into the upper sign of (7.214), we obtain

$$\left|\frac{3}{2},\, \frac{3}{2}\right\rangle = \left|1, \frac{1}{2};\, 1, \frac{1}{2}\right\rangle, \tag{7.215}$$

$$\left|\frac{3}{2},\, \frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\left|1, \frac{1}{2};\, 0, \frac{1}{2}\right\rangle + \frac{1}{\sqrt{3}}\left|1, \frac{1}{2};\, 1, -\frac{1}{2}\right\rangle, \tag{7.216}$$

$$\left|\frac{3}{2},\, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\left|1, \frac{1}{2};\, -1, \frac{1}{2}\right\rangle + \sqrt{\frac{2}{3}}\left|1, \frac{1}{2};\, 0, -\frac{1}{2}\right\rangle, \tag{7.217}$$

$$\left|\frac{3}{2},\, -\frac{3}{2}\right\rangle = \left|1, \frac{1}{2};\, -1, -\frac{1}{2}\right\rangle. \tag{7.218}$$

Similarly, an insertion of $l = 1$ and $m = \frac{1}{2}, -\frac{1}{2}$ into the lower sign of (7.214) yields

$$\left|\frac{1}{2},\, \frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\left|1, \frac{1}{2};\, 1, -\frac{1}{2}\right\rangle - \frac{1}{\sqrt{3}}\left|1, \frac{1}{2};\, 0, \frac{1}{2}\right\rangle, \tag{7.219}$$

$$\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\left|1, \frac{1}{2};\, 0, -\frac{1}{2}\right\rangle - \sqrt{\frac{2}{3}}\left|1, \frac{1}{2};\, -1, \frac{1}{2}\right\rangle. \tag{7.220}$$

Spin–orbit functions

The eigenfunctions of the particle's total angular momentum $\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$ may be represented by the direct product of the eigenstates of the orbital and spin angular momenta, $\left|l, m - \frac{1}{2}\right\rangle$ and $\left|\frac{1}{2}, \frac{1}{2}\right\rangle$. From (7.214) we have

$$\left|l \pm \frac{1}{2}, m\right\rangle = \sqrt{\frac{l \mp m + \frac{1}{2}}{2l+1}}\left|l, m + \frac{1}{2}\right\rangle\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \pm \sqrt{\frac{l \pm m + \frac{1}{2}}{2l+1}}\left|l, m - \frac{1}{2}\right\rangle\left|\frac{1}{2}, \frac{1}{2}\right\rangle. \tag{7.221}$$

If this particle moves in a central potential, its complete wave function consists of a space part, $\langle r\theta\varphi | n, l, m \pm \frac{1}{2}\rangle = R_{nl}(r)Y_{l,m\pm\frac{1}{2}}$, and a spin part, $\left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle$:

$$\Psi_{n,l,j=l\pm\frac{1}{2},m} = R_{nl}(r)\left[\sqrt{\frac{l \mp m + \frac{1}{2}}{2l+1}}Y_{l,m+\frac{1}{2}}\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \pm \sqrt{\frac{l \pm m + \frac{1}{2}}{2l+1}}Y_{l,m-\frac{1}{2}}\left|\frac{1}{2}, \frac{1}{2}\right\rangle\right], \tag{7.222}$$

Using the spinor representation for the spin part, $\left|\frac{1}{2}, \frac{1}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, we can

write (7.222) as follows:

$$\Psi_{n,l,j=l\pm\frac{1}{2},m}(r,\theta,\varphi) = \frac{R_{nl}(r)}{\sqrt{2l+1}} \begin{pmatrix} \pm\sqrt{l\pm m+\frac{1}{2}}\,Y_{l,m-\frac{1}{2}}(\theta,\varphi) \\ \sqrt{l\mp m+\frac{1}{2}}\,Y_{l,m+\frac{1}{2}}(\theta,\varphi) \end{pmatrix}, \tag{7.223}$$

where m is half-integer. The states (7.222) and (7.223) are simultaneous eigenfunctions of $\hat{\vec{J}}^2$, $\hat{\vec{L}}^2$, $\hat{\vec{S}}^2$, and $\hat{J}_z$ with eigenvalues $\hbar^2 j(j+1)$, $\hbar^2 l(l+1)$, $\hbar^2 s(s+1) = 3\hbar^2/4$, and $\hbar m$, respectively. The wave functions $\Psi_{n,l,j=l\pm\frac{1}{2},m}(r,\theta,\varphi)$ are eigenstates of $\hat{\vec{L}}\cdot\hat{\vec{S}}$ as well, since

$$\begin{aligned} \hat{\vec{L}}\cdot\hat{\vec{S}}\,|nljm\rangle &= \frac{1}{2}\left(\hat{\vec{J}}^2 - \hat{\vec{L}}^2 - \hat{\vec{S}}^2\right)|nljm\rangle \\ &= \frac{\hbar^2}{2}\left[j(j+1) - l(l+1) - s(s+1)\right]|nljm\rangle\,. \end{aligned} \tag{7.224}$$

Here j takes only two values, $j = l \pm \frac{1}{2}$, so we have

$$\langle nljm|\,\hat{\vec{L}}\cdot\hat{\vec{S}}\,|nljm\rangle = \frac{\hbar^2}{2}\left[j(j+1) - l(l+1) - \frac{3}{4}\right] = \begin{cases} \frac{1}{2}l\hbar^2, & j = l+\frac{1}{2}, \\ -\frac{1}{2}(l+1)\hbar^2, & j = l-\frac{1}{2}. \end{cases} \tag{7.225}$$

7.3.4 Addition of More Than Two Angular Momenta

The formalism for adding two angular momenta may be generalized to those cases where we add three or more angular momenta. For instance, to add three mutually commuting angular momenta $\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2 + \hat{\vec{J}}_3$, we may follow any of these three methods. (a) Add $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ to obtain $\hat{\vec{J}}_{12} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2$, and then add $\hat{\vec{J}}_{12}$ to $\hat{\vec{J}}_3$: $\hat{\vec{J}} = \hat{\vec{J}}_{12} + \hat{\vec{J}}_3$. (b) Add $\hat{\vec{J}}_2$ and $\hat{\vec{J}}_3$ to form $\hat{\vec{J}}_{23} = \hat{\vec{J}}_2 + \hat{\vec{J}}_3$, and then add $\hat{\vec{J}}_{23}$ to $\hat{\vec{J}}_1$: $\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_{23}$. (c) Add $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_3$ to form $\hat{\vec{J}}_{13} = \hat{\vec{J}}_1 + \hat{\vec{J}}_3$, and then add $\hat{\vec{J}}_{13}$ to $\hat{\vec{J}}_2$: $\hat{\vec{J}} = \hat{\vec{J}}_2 + \hat{\vec{J}}_{13}$.

Considering the first method and denoting the eigenstates of $\hat{\vec{J}}_1^2$ and $\hat{J}_{1z}$ by $|\,j_1,\,m_1\rangle$, those of $\hat{\vec{J}}_2^2$ and $\hat{J}_{2z}$ by $|\,j_2,\,m_2\rangle$, and those of $\hat{\vec{J}}_3^2$ and $\hat{J}_{3z}$ by $|\,j_3,\,m_3\rangle$, we may express the joint eigenstates $|j_{12},j,m\rangle$ of $\hat{\vec{J}}_1^2$, $\hat{\vec{J}}_2^2$, $\hat{\vec{J}}_3^2$, $\hat{\vec{J}}_{12}^2$, $\hat{\vec{J}}^2$ and J_z in terms of the states

$$|j_1,j_2,j_3;\,m_1,m_2,m_3\rangle = |\,j_1,\,m_1\rangle\,|\,j_2,\,m_2\rangle|j_3,\,m_3\rangle \tag{7.226}$$

as follows. First, the coupling of $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ leads to

$$|j_{12},\,m_{12}\rangle = \sum_{m_1=-j_1}^{j_1}\sum_{m_2=-j_2}^{j_2}\langle j_1,j_2;\,m_1,m_2\,|\,j_{12},\,m_{12}\rangle|j_1,j_2;\,m_1,m_2\rangle, \tag{7.227}$$

where $m_{12} = m_1 + m_2$ and $|j_1 - j_2| \le j_{12} \le |j_1 + j_2|$. Then, adding $\hat{\vec{J}}_{12}$ and $\hat{\vec{J}}_3$, the state $|j_{12},j,m\rangle$ is given by

$$\sum_{m_{12}=-j_{12}}^{j_{12}}\sum_{m_3=-j_3}^{j_3}\langle j_1,j_2;m_1,m_2\,|\,j_{12},m_{12}\rangle\langle j_{12},j_3;m_{12},m_3|j_{12},j,m\rangle|j_1,j_2,j_3;m_1,m_2,m_3\rangle, \tag{7.228}$$

with $m = m_{12}+m_3$ and $|j_{12}-j_3| \le j \le |j_{12}+j_3|$; the Clebsch–Gordan coefficients $\langle j_1, j_2;\ m_1, m_2 | j_{12},\ m_{12}\rangle$ and $\langle j_{12}, j_3;\ m_{12}, m_3 | j_{12}, j, m\rangle$ correspond to the coupling of $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ and of $\hat{\vec{J}}_{12}$ and $\hat{\vec{J}}_3$, respectively. The calculation of these coefficients is similar to that of two angular momenta. For instance, in Problem 7.4, page 484, we will see how to add three spins and how to calculate the corresponding Clebsch–Gordan coefficients.

We should note that the addition of $\hat{\vec{J}}_1$, $\hat{\vec{J}}_2$, and $\hat{\vec{J}}_3$ in essence consists of constructing the eigenvectors $|j_{12}, j, m\rangle$ in terms of the $(2j_1+1)(2j_2+1)(2j_3+1)$ states $|\ j_1, j_2, j_3;\ m_1, m_2, m_3\rangle$. We may then write

$$\hat{J}_{\pm}|j_{12}, j, m\rangle = \hbar\sqrt{j(j+1) - m(m\pm 1)}|j_{12}, j, m\pm 1\rangle, \tag{7.229}$$

$$\hat{J}_{1_\pm}\ |\ j_1, j_2, j_3;\ m_1, m_2, m_3\rangle = \hbar\sqrt{j_1(j_1+1) - m_1(m_1\pm 1)}|j_1, j_2, j_3;\ (m_1\pm 1), m_2, m_3\rangle, \tag{7.230}$$

$$\hat{J}_{2_\pm}\ |\ j_1, j_2, j_3;\ m_1, m_2, m_3\rangle = \hbar\sqrt{j_2(j_2+1) - m_2(m_2\pm 1)}|j_1, j_2, j_3;\ m_1, (m_2\pm 1), m_3\rangle, \tag{7.231}$$

$$\hat{J}_{3_\pm}\ |\ j_1, j_2, j_3;\ m_1, m_2, m_3\rangle = \hbar\sqrt{j_3(j_3+1) - m_3(m_3\pm 1)}|j_1, j_2, j_3;\ m_1, m_2, (m_3\pm 1)\rangle. \tag{7.232}$$

The foregoing method can be generalized to the coupling of more than three angular momenta: $\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2 + \hat{\vec{J}}_3 + \cdots + \hat{\vec{J}}_N$. Each time we couple two angular momenta, we reduce the problem to the coupling of $(N-1)$ angular momenta. For instance, we may start by adding $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ to generate $\hat{\vec{J}}_{12}$; we are then left with $(N-1)$ angular momenta. Second, by adding $\hat{\vec{J}}_{12}$ and $\hat{\vec{J}}_3$ to form $\hat{\vec{J}}_{123}$, we are left with $(N-2)$ angular momenta. Third, an addition of $\hat{\vec{J}}_{123}$ and $\hat{\vec{J}}_4$ leaves us with $(N-3)$ angular momenta, and so on. We may continue in this way till we add all given angular momenta.

7.3.5 Rotation Matrices for Coupling Two Angular Momenta

We want to find out how to express the rotation matrix associated with an angular momentum $\hat{\vec{J}}$ in terms of the rotation matrices corresponding to $\hat{\vec{J}}_1$ and $\hat{\vec{J}}_2$ such that $\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2$. That is, knowing the rotation matrices $d^{(j_1)}(\beta)$ and $d^{(j_2)}(\beta)$, how does one calculate $d^{(j)}_{mm'}(\beta)$?

Since

$$d^{(j)}_{m'm}(\beta) = \langle j,\ m' \mid \hat{R}_y(\beta) \mid j,\ m\rangle, \tag{7.233}$$

where

$$|\ j,\ m\rangle = \sum_{m_1 m_2}\langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle \mid j_1, j_2;\ m_1, m_2\rangle, \tag{7.234}$$

$$|\ j,\ m'\rangle = \sum_{m_1' m_2'}\langle j_1, j_2;\ m_1', m_2' \mid j,\ m'\rangle|j_1, j_2;\ m_1', m_2'\rangle, \tag{7.235}$$

and since the Clebsch–Gordan coefficients are real,

$$\langle j,\ m'| = \sum_{m_1' m_2'}\langle j_1, j_2;\ m_1', m_2' \mid j,\ m'\rangle\langle j_1, j_2;\ m_1', m_2'|, \tag{7.236}$$

we can rewrite (7.233) as

$$d^{(j)}_{m'm}(\beta) = \sum_{m_1 m_2} \sum_{m'_1 m'_2} \langle j_1, j_2;\, m_1, m_2 \mid j,\, m\rangle \langle j_1, j_2;\, m'_1, m'_2 \mid j,\, m'\rangle \times \langle j_1, j_2;\, m'_1, m'_2 | \hat{R}_y(\beta) \mid j_1, j_2;\, m_1, m_2\rangle. \tag{7.237}$$

Since $\hat{R}_y(\beta) = \exp[-i\beta \hat{J}_y/\hbar] = \exp[-i\beta \hat{J}_{1_y}/\hbar]\exp[-i\beta \hat{J}_{2_y}/\hbar]$, because $\hat{J}_y = \hat{J}_{1_y} + \hat{J}_{2_y}$, and since $\langle j_1, j_2;\, m'_1, m'_2| = \langle j_1,\, m'_1 \mid \langle j_2,\, m'_2 \mid$ and $\mid j_1, j_2;\, m_1, m_2\rangle = \mid j_1,\, m_1\rangle \mid j_2,\, m_2\rangle$, we have

$$d^{(j)}_{m'm}(\beta) = \sum_{m_1 m_2} \sum_{m'_1 m'_2} \langle j_1, j_2;\, m_1, m_2 \mid j,\, m\rangle \langle j_1, j_2;\, m'_1, m'_2 \mid j,\, m'\rangle \times \langle j_1,\, m'_1 \mid \exp\left[-\frac{i}{\hbar}\beta \hat{J}_{1_y}\right] \mid j_1,\, m_1\rangle \langle j_2,\, m'_2 \mid \exp\left[-\frac{i}{\hbar}\beta \hat{J}_{2_y}\right] \mid j_2,\, m_2\rangle, \tag{7.238}$$

or

$$\boxed{d^{(j)}_{m'm}(\beta) = \sum_{m_1 m_2} \sum_{m'_1 m'_2} \langle j_1, j_2;\, m_1, m_2 \mid j,\, m\rangle \langle j_1, j_2;\, m'_1, m'_2 \mid j,\, m'\rangle d^{(j_1)}_{m'_1 m_1}(\beta) d^{(j_2)}_{m'_2 m_2}(\beta),} \tag{7.239}$$

with

$$d^{(j_1)}_{m'_1 m_1}(\beta) = \langle j_1,\, m'_1 \mid \exp\left[-\frac{i}{\hbar}\beta \hat{J}_{1_y}\right] \mid j_1,\, m_1\rangle, \tag{7.240}$$

$$d^{(j_2)}_{m'_2 m_2}(\beta) = \langle j_2,\, m'_2 \mid \exp\left[-\frac{i}{\hbar}\beta \hat{J}_{2_y}\right] \mid j_2,\, m_2\rangle. \tag{7.241}$$

From (7.62) we have

$$d^{(j)}_{m'm}(\beta) = e^{i(m'\alpha + m\gamma)} D^{(j)}_{m'm}(\alpha, \beta, \gamma); \tag{7.242}$$

hence can rewrite (7.239) as

$$\boxed{D^{(j)}_{m'm}(\alpha, \beta, \gamma) = \sum_{m_1 m_2} \sum_{m'_1 m'_2} \langle j_1, j_2;\, m_1, m_2 \mid j,\, m\rangle \langle j_1, j_2;\, m'_1, m'_2 \mid j,\, m'\rangle D^{(j_1)}_{m'_1 m_1}(\alpha, \beta, \gamma) D^{(j_2)}_{m'_2 m_2}(\alpha, \beta, \gamma),} \tag{7.243}$$

since $m = m_1 + m_2$ and $m' = m'_1 + m'_2$.

Now, let us see how to express the product of the rotation matrices $d^{(j_1)}(\beta)$ and $d^{(j_2)}(\beta)$ in terms of $d^{(j)}_{mm'}(\beta)$. Sandwiching both sides of

$$\exp\left[-\frac{i}{\hbar}\beta \hat{J}_{1_y}\right] \exp\left[-\frac{i}{\hbar}\beta \hat{J}_{2_y}\right] = \exp\left[-\frac{i}{\hbar}\beta \hat{J}_y\right] \tag{7.244}$$

between

$$\mid j_1, j_2;\, m_1, m_2\rangle = \sum_{jm} \langle j_1, j_2;\, m_1, m_2 \mid j,\, m\rangle \mid j,\, m\rangle \tag{7.245}$$

and

$$\langle j_1, j_2;\, m'_1, m'_2| = \sum_{jm'} \langle j_1, j_2;\, m'_1, m'_2 \mid j,\, m'\rangle \langle j,\, m' \mid, \tag{7.246}$$

and since $\langle j_1, j_2;\ m_1', m_2'| = \langle j_1,\ m_1' \mid \langle j_2,\ m_2' \mid$ and $\mid j_1, j_2;\ m_1, m_2\rangle = \mid j_1,\ m_1\rangle \mid j_2,\ m_2\rangle$, we have

$$\langle j_1,\ m_1' \mid \exp\left[-\frac{i}{\hbar}\beta \hat{J}_{1_y}\right] \mid j_1,\ m_1\rangle\langle j_2,\ m_2' \mid \exp\left[-\frac{i}{\hbar}\beta \hat{J}_{2_y}\right] \mid j_2,\ m_2\rangle$$
$$= \sum_{jmm'} \langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle\langle j_1, j_2;\ m_1', m_2' \mid j,\ m'\rangle\langle j,\ m' \mid \hat{R}_y(\beta) \mid j,\ m\rangle \tag{7.247}$$

or

$$\boxed{d^{(j_1)}_{m_1' m_1}(\beta) d^{(j_2)}_{m_2' m_2}(\beta) = \sum_{|j_1 - j_2|}^{|j_1 + j_2|} \sum_{mm'} \langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle\langle j_1, j_2;\ m_1', m_2' \mid j,\ m'\rangle d^{(j)}_{m' m}(\beta).} \tag{7.248}$$

Following the same procedure that led to (7.243), we can rewrite (7.248) as

$$\boxed{D^{(j_1)}_{m_1' m_1}(\alpha, \beta, \gamma) D^{(j_2)}_{m_2' m_2}(\alpha, \beta, \gamma) = \sum_{jmm'} \langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle\langle j_1, j_2;\ m_1', m_2' \mid j,\ m'\rangle D^{(j)}_{m' m}(\alpha, \beta, \gamma).} \tag{7.249}$$

This relation is known as the *Clebsch–Gordan series*.

The relation (7.249) has an important application: the derivation of an integral involving three spherical harmonics. When j_1 and j_2 are both integers (i.e., $j_1 = l_1$ and $j_2 = l_2$) and m_1 and m_2 are both zero (hence $m = 0$), equation (7.249) finds a useful application:

$$D^{(l_1)}_{m_1' 0}(\alpha, \beta, \gamma) D^{(l_2)}_{m_2' 0}(\alpha, \beta, \gamma) = \sum_{lm'} \langle l_1, l_2;\ 0, 0 \mid l,\ 0\rangle\langle l_1, l_2;\ m_1', m_2' \mid l,\ m'\rangle D^{(l)}_{m' 0}(\alpha, \beta, \gamma). \tag{7.250}$$

Since the expressions of $D^{(l_1)}_{m_1' 0}$, $D^{(l_2)}_{m_2' 0}$, and $D^{(l)}_{m' 0}$ can be inferred from (7.81), notably

$$D^{(l)}_{m' 0}(\alpha, \beta, 0) = \sqrt{\frac{4\pi}{2l + 1}} Y_{lm'}(\beta, \alpha), \tag{7.251}$$

we can reduce (7.250) to

$$\boxed{Y_{l_1 m_1}(\beta, \alpha) Y_{l_2 m_2}(\beta, \alpha) = \sum_{lm} \sqrt{\frac{(2l_1 + 1)(2l_2 + 1)}{4\pi(2l + 1)}} \langle l_1, l_2;\ 0, 0 \mid l,\ 0\rangle\langle l_1, l_2;\ m_1, m_2 \mid l,\ m\rangle Y_{lm}(\beta, \alpha),} \tag{7.252}$$

where we have removed the primes and taken the complex conjugate. Multiplying both sides by $Y^*_{lm}(\beta, \alpha)$ and integrating over α and β, we obtain the following frequently used integral:

$$\boxed{\int_0^{2\pi} d\alpha \int_0^{\pi} Y^*_{lm}(\beta, \alpha) Y_{l_1 m_1}(\beta, \alpha) Y_{l_2 m_2}(\beta, \alpha) \sin\beta\, d\beta = \sqrt{\frac{(2l_1 + 1)(2l_2 + 1)}{4\pi(2l + 1)}} \langle l_1, l_2; 0, 0 \mid l,\ 0\rangle \times \langle l_1, l_2;\ m_1, m_2 \mid l,\ m\rangle.} \tag{7.253}$$

7.3.6 Isospin

The ideas presented above—spin and the addition of angular momenta—find some interesting applications to other physical quantities. For instance, in the field of nuclear physics, the quantity known as *isotopic spin* can be represented by a set of operators which not only obey the same algebra as the components of angular momentum, but also couple in the same way as ordinary angular momenta.

Since the nuclear force does not depend on the electric charge, we can consider the proton and the neutron to be separate manifestations (states) of the same particle, the *nucleon*. The nucleon may thus be found in two different states: a proton and a neutron. In this way, as the protons and neutrons are identical particles with respect to the nuclear force, we will need an additional quantum number (or label) to indicate whether the nucleon is a proton or a neutron. Due to its formal analogy with ordinary spin, this label is called the *isotopic spin* or, in short, the *isospin*. If we take the isospin quantum number to be $\frac{1}{2}$, its z-component will then be represented by a quantum number having the values $\frac{1}{2}$ and $-\frac{1}{2}$. The difference between a proton and a neutron then becomes analogous to the difference between spin-up and spin-down particles.

The fundamental difference between ordinary spin and the isospin is that, unlike the spin, the isospin has nothing to do with rotations or spinning in the coordinate space, it hence cannot be coupled with the angular momenta of the nucleons. Nucleons can thus be distinguished by $\langle \hat{t}_3 \rangle = \pm\frac{1}{2}$, where $\hat{t}_3$ is the third or z-component of the isospin vector operator $\hat{\vec{t}}$.

7.3.6.1 Isospin Algebra

Due to the formal analogy between the isospin and the spin, their formalisms have similar structures from a mathematical viewpoint. The algebra obeyed by the components $\hat{t}_1, \hat{t}_2, \hat{t}_3$ of the isospin operator $\hat{\vec{t}}$ can thus be inferred from the properties and commutation relations of the spin operator. For instance, the components of the isospin operator can be constructed from the Pauli matrices $\vec{\tau}$ in the same way as we did for the angular momentum operators of spin $\frac{1}{2}$ particles:

$$\hat{\vec{t}} = \frac{1}{2}\vec{\tau}, \tag{7.254}$$

with

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{7.255}$$

The components $\hat{t}_1, \hat{t}_2, \hat{t}_3$ obey the same commutation relations as those of angular momentum:

$$\boxed{\left[\hat{t}_1, \hat{t}_2\right] = i\hat{t}_3, \qquad \left[\hat{t}_2, \hat{t}_3\right] = i\hat{t}_1, \qquad \left[\hat{t}_3, \hat{t}_1\right] = i\hat{t}_2.} \tag{7.256}$$

So the nucleon can be found in two different states: when $\hat{t}_3$ acts on a nucleon state, it gives the eignvalues $\pm\frac{1}{2}$. By convention the $\hat{t}_3$ of a proton is taken to be $\hat{t}_3 = +\frac{1}{2}$ and that of a neutron is $\hat{t}_3 = -\frac{1}{2}$. Denoting the proton and neutron states, respectively, by $|\, p\rangle$ and $|\, n\rangle$,

$$|\, p\rangle = \left|t = \frac{1}{2}, t_3 = \frac{1}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |\, n\rangle = \left|t = \frac{1}{2}, t_3 = -\frac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{7.257}$$

we have

$$\hat{t}_3 \mid p\rangle = \hat{t}_3 \left|\frac{1}{2}, \frac{1}{2}\right\rangle = \frac{1}{2}\left|\frac{1}{2}, \frac{1}{2}\right\rangle, \tag{7.258}$$

$$\hat{t}_3 \mid n\rangle = \hat{t}_3 \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = -\frac{1}{2}\left|\frac{1}{2}, -\frac{1}{2}\right\rangle. \tag{7.259}$$

We can write (7.258) and (7.259), respectively, as

$$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{7.260}$$

$$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{7.261}$$

By analogy with angular momentum, denoting the joint eigenstates of $\hat{\vec{t}}^2$ and $\hat{t}_3$ by $\mid t,\ t_3\rangle$, we have

$$\boxed{\hat{\vec{t}}^2 \mid t,\ t_3\rangle = t(t+1) \mid t,\ t_3\rangle, \qquad \hat{t}_3 \mid t,\ t_3\rangle = t_3 \mid t,\ t_3\rangle.} \tag{7.262}$$

We can also introduce the *raising and lowering* isospin operators:

$$\hat{t}_+ = \hat{t}_1 + i\hat{t}_2 = \frac{1}{2}(\tau_1 + i\tau_2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{7.263}$$

$$\hat{t}_- = \hat{t}_1 - i\hat{t}_2 = \frac{1}{2}(\tau_1 - i\tau_2) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{7.264}$$

hence

$$\boxed{\hat{t}_\pm \mid t,\ t_3\rangle = \sqrt{t(t+1) - t_3(t_3 \pm 1)}\ \mid t,\ t_3 \pm 1\rangle.} \tag{7.265}$$

Note that $\hat{t}_+$ and $\hat{t}_-$ are operators which, when acting on a nucleon state, convert neutron states into proton states and proton states into neutron states, respectively:

$$\hat{t}_+ \mid n\rangle = \mid p\rangle, \qquad \hat{t}_- \mid p\rangle = \mid n\rangle. \tag{7.266}$$

We can also define a *charge operator*

$$\hat{Q} = e\left(\hat{t}_3 + \frac{1}{2}\right), \tag{7.267}$$

where e is the charge of the proton, with

$$\hat{Q} \mid p\rangle = e \mid p\rangle, \qquad \hat{Q} \mid n\rangle = 0. \tag{7.268}$$

We should mention that strong interactions conserve isospin. For instance, a reaction like

$$d + d \rightarrow \alpha + \pi^0 \tag{7.269}$$

is forbidden since the isospin is not conserved, because the isospins of d and α are both zero and the isospin of the pion is equal to one (i.e., $T(d) = T(\alpha) = 0$, but $T(\pi) = 1$); this leads to isospin zero for $(d+d)$ and isospin one for $(\alpha+\pi^0)$. The reaction was confirmed experimentally to be forbidden, since its cross-section is negligibly small. However, reactions such as

$$p + p \rightarrow d + \pi^+, \quad p + n \rightarrow d + \pi^0 \tag{7.270}$$

are allowed, since they conserve isospin.

7.3.6.2 Addition of Two Isospins

We should note that the isospins of different nucleons can be added in the same way as adding angular momenta. For a nucleus consisting of several nucleons, the total isospin is given by the vector sum of the isospins of all individual nucleons: $\hat{\vec{T}} = \sum_i^A \hat{\vec{t}}_i$. For instance, the total isospin of a system of two nucleons can be obtained by coupling their isospins $\hat{\vec{t}}_1$ and $\hat{\vec{t}}_2$:

$$\hat{\vec{T}} = \hat{\vec{t}}_1 + \hat{\vec{t}}_2. \tag{7.271}$$

Denoting the joint eigenstates of $\hat{\vec{t}}_1^2$, $\hat{\vec{t}}_2^2$, $\hat{\vec{T}}^2$, and $\hat{T}_3$ by $|\, T,\, N\rangle$, we have:

$$\hat{\vec{T}}^2 \,|\, T,\, N\rangle = T(T+1)\,|\, T,\, N\rangle, \qquad \hat{T}_3 \,|\, T,\, N\rangle = N \,|\, T,\, N\rangle. \tag{7.272}$$

Similarly, if we denote the joint eigenstates of $\hat{\vec{t}}_1^2$, $\hat{\vec{t}}_2^2$, $\hat{t}_{1_3}$, and $\hat{t}_{2_3}$ by $|t_1, t_2;\, n_1, n_2\rangle$, we have

$$\hat{\vec{t}}_1^2 |t_1, t_2;\, n_1, n_2\rangle = t_1(t_1+1)|t_1, t_2;\, n_1, n_2\rangle, \tag{7.273}$$

$$\hat{\vec{t}}_2^2 |t_1, t_2;\, n_1, n_2\rangle = t_2(t_2+1)|t_1, t_2;\, n_1, n_2\rangle, \tag{7.274}$$

$$\hat{t}_{1_3} |t_1, t_2;\, n_1, n_2\rangle = n_1 |t_1, t_2;\, n_1, n_2\rangle, \tag{7.275}$$

$$\hat{t}_{2_3} |t_1, t_2;\, n_1, n_2\rangle = n_2 |t_1, t_2;\, n_1, n_2\rangle, \tag{7.276}$$

The matrix elements of the unitary transformation connecting the $\{|T,\, N\rangle\}$ and $\{|t_1, t_2;\, n_1, n_2\rangle\}$ bases,

$$|T,\, N\rangle = \sum_{n_1, n_2} \langle t_1, t_2;\, n_1 n_2 | T,\, N\rangle |t_1, t_2;\, n_1, n_2\rangle, \tag{7.277}$$

are given by the coefficients $\langle t_1, t_2;\, n_1 n_2 | T,\, N\rangle$; these coefficients can be calculated in the same way as the Clebsch–Gordan coefficients; see the next example.

Example 7.4
Find the various states corresponding to a two-nucleon system.

Solution
Let $\hat{\vec{T}}$ be the total isospin vector operator of the two-nucleon system:

$$\hat{\vec{T}} = \hat{\vec{t}}_1 + \hat{\vec{t}}_2. \tag{7.278}$$

This example is similar to adding two spin $\frac{1}{2}$ angular momenta. Thus, the values of T are 0 and 1. The case $T = 0$ corresponds to a singlet state:

$$|\, 0,\, 0\rangle = \frac{1}{\sqrt{2}} \left[|\, p\rangle_1 \,|\, n\rangle_2 - |\, n\rangle_1 \,|\, p\rangle_2 \right], \tag{7.279}$$

where $|\, p\rangle_1$ means that nucleon 1 is a proton, $|\, n\rangle_2$ means that nucleon 2 is a neutron, and so on. This state, which is an antisymmetric isospin state, describes a bound (p-n) system such as the ground state of deuterium ($T = 0$).

The case $T = 1$ corresponds to the triplet states $|1, N\rangle$ with $N = 1, 0, -1$:

$$|1, 1\rangle = |p\rangle_1 |p\rangle_2, \tag{7.280}$$

$$|1, 0\rangle = \frac{1}{\sqrt{2}} [|p\rangle_1 |n\rangle_2 + |n\rangle_1 |p\rangle_2], \tag{7.281}$$

$$|1, -1\rangle = |n\rangle_1 |n\rangle_2. \tag{7.282}$$

The state $|1, 1\rangle$ corresponds to the case where both nucleons are protons (p-p) and $|1, -1\rangle$ corresponds to the case where both nucleons are neutrons (n-n).

7.4 Rotations of Scalar, Vector, and Tensor Operators

In this section, we are going to study how operators transform under rotations. Operators corresponding to various physical quantities can be classified as scalars, vectors, and tensors, depending on their behavior under rotations.

Consider an operator $\hat{A}$, which can be a scalar, a vector, or a tensor. In accordance with (7.23), the transformation of $\hat{A}$ under a rotation with an infinitesimal angle of $\delta\theta$ about an axis $\vec{n}$ is given by

$$\hat{A}' = \hat{R}_n(\delta\theta)\, \hat{A}\, \hat{R}_n^\dagger(\delta\theta), \tag{7.283}$$

where $\hat{R}_n(\delta\theta)$ can be inferred from (7.28)

$$\hat{R}_n(\delta\theta) = 1 - \frac{i}{\hbar}\delta\theta \vec{n} \cdot \hat{\vec{J}}, \tag{7.284}$$

where $\hat{\vec{J}}$ is the total angular momentum operator. Substituting (7.284) into (7.283) and keeping terms up to first order in $\delta\theta$, we obtain

$$\boxed{\hat{A}' = \hat{A} + \frac{i}{\hbar}\delta\theta [\hat{A}, \vec{n} \cdot \hat{\vec{J}}].} \tag{7.285}$$

In the rest of this section we focus on the application of this relation to scalar, vector, and tensor operators.

7.4.1 Transformation of Scalar Operators Under Rotations

Since scalar operators are invariant under rotations (i.e., $\hat{A}' = \hat{A}$), equation (7.285) implies that they commute with the angular momentum

$$\boxed{[\hat{A}, \hat{J}_k] = 0 \qquad (k = x, y, z).} \tag{7.286}$$

This is also true for pseudo-scalars. A pseudo-scalar is defined by the product of a vector $\vec{A}$ and a pseudo-vector or axial vector $\vec{B} \times \vec{C}$: $\vec{A} \cdot (\vec{B} \times \vec{C})$.

7.4.2 Transformation of Vector Operators Under Rotations

On the one hand, a vector operator $\hat{\vec{A}}$ transforms according to (7.285):

$$\boxed{\hat{\vec{A}}' = \hat{\vec{A}} + \frac{i}{\hbar}\delta\theta\,[\hat{\vec{A}},\ \vec{n}\cdot\hat{\vec{J}}\,].} \tag{7.287}$$

On the other hand, from the classical theory of rotations, when a vector $\hat{\vec{A}}$ is rotated through an angle $\delta\theta$ around an axis $\vec{n}$, it is given by

$$\hat{\vec{A}}' = \hat{\vec{A}} - \delta\theta\,\vec{n}\times\hat{\vec{A}}. \tag{7.288}$$

Comparing (7.287) and (7.288), we obtain

$$\boxed{[\hat{\vec{A}},\ \vec{n}\cdot\hat{\vec{J}}] = i\hbar\vec{n}\times\hat{\vec{A}}.} \tag{7.289}$$

The jth component of this equation is given by

$$[\hat{\vec{A}},\ \vec{n}\cdot\hat{\vec{J}}]_j = i\hbar(\vec{n}\times\hat{\vec{A}})_j \qquad (j = x, y, z), \tag{7.290}$$

which in the case of $j = x, y, z$ leads to

$$\left[\hat{A}_x, \hat{J}_x\right] = \left[\hat{A}_y, \hat{J}_y\right] = \left[\hat{A}_z, \hat{J}_z\right] = 0, \tag{7.291}$$

$$\left[\hat{A}_x, \hat{J}_y\right] = i\hbar\hat{A}_z, \quad \left[\hat{A}_y, \hat{J}_z\right] = i\hbar\hat{A}_x, \quad \left[\hat{A}_z, \hat{J}_x\right] = i\hbar\hat{A}_y, \tag{7.292}$$

$$\left[\hat{A}_x, \hat{J}_z\right] = -i\hbar\hat{A}_y, \quad \left[\hat{A}_y, \hat{J}_x\right] = -i\hbar\hat{A}_z, \quad \left[\hat{A}_z, \hat{J}_y\right] = -i\hbar\hat{A}_x. \tag{7.293}$$

Some interesting applications of (7.289) correspond to the cases where the vector operator $\hat{\vec{A}}$ is either the angular momentum, the position, or the linear momentum operator. Let us consider these three cases separately. First, substituting $\hat{\vec{A}} = \hat{\vec{J}}$ into (7.289), we recover the usual angular momentum commutation relations:

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z, \qquad [\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x, \qquad [\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y. \tag{7.294}$$

Second, in the case of a spinless particle (i.e., $\hat{\vec{J}} = \hat{\vec{L}}$), and if $\hat{\vec{A}}$ is equal to the position operator, $\hat{\vec{A}} = \vec{R}$, then (7.289) will yield the following relations:

$$\left[\hat{x}, \hat{L}_x\right] = 0, \quad \left[\hat{x}, \hat{L}_y\right] = i\hbar\hat{z}, \quad \left[\hat{x}, \hat{L}_z\right] = -i\hbar\hat{y}, \tag{7.295}$$

$$\left[\hat{y}, \hat{L}_y\right] = 0, \quad \left[\hat{y}, \hat{L}_z\right] = i\hbar\hat{x}, \quad \left[\hat{y}, \hat{L}_x\right] = -i\hbar\hat{z}, \tag{7.296}$$

$$\left[\hat{z}, \hat{L}_z\right] = 0, \quad \left[\hat{z}, \hat{L}_x\right] = i\hbar\hat{y}, \quad \left[\hat{z}, \hat{L}_y\right] = -i\hbar\hat{x}. \tag{7.297}$$

Third, if $\hat{\vec{J}} = \hat{\vec{L}}$ and if $\hat{\vec{A}}$ is equal to the momentum operator, $\hat{\vec{A}} = \hat{\vec{P}}$, then (7.289) will lead to

$$\left[\hat{P}_x, \hat{L}_x\right] = 0, \quad \left[\hat{P}_x, \hat{L}_y\right] = i\hbar\hat{P}_z, \quad \left[\hat{P}_x, \hat{L}_z\right] = -i\hbar\hat{P}_y, \tag{7.298}$$

$$\left[\hat{P}_y, \hat{L}_y\right] = 0, \quad \left[\hat{P}_y, \hat{L}_z\right] = i\hbar\hat{P}_x, \quad \left[\hat{P}_y, \hat{L}_x\right] = -i\hbar\hat{P}_z, \tag{7.299}$$

$$\left[\hat{P}_z, \hat{L}_z\right] = 0, \quad \left[\hat{P}_z, \hat{L}_x\right] = i\hbar\hat{P}_y, \quad \left[\hat{P}_z, \hat{L}_y\right] = -i\hbar\hat{P}_x. \tag{7.300}$$

Now, introducing the operators

$$\hat{A}_{\pm} = \hat{A}_x \pm i\hat{A}_y, \tag{7.301}$$

and using the relations (7.291) to (7.293), we can show that

$$\left[\hat{J}_x, \hat{A}_{\pm}\right] = \mp\hbar\hat{A}_z, \qquad \left[J_y, \hat{A}_{\pm}\right] = -i\hbar\hat{A}_z, \qquad \left[J_z, \hat{A}_{\pm}\right] = \pm\hbar\hat{A}_{\pm}. \tag{7.302}$$

These relations in turn can be shown to lead to

$$\left[\hat{J}_{\pm}, \hat{A}_{\pm}\right] = 0, \qquad \left[\hat{J}_{\pm}, \hat{A}_{\mp}\right] = \pm 2\hbar\hat{A}_z. \tag{7.303}$$

Let us introduce the *spherical components* $\hat{A}_{-1}$, $\hat{A}_0$, $\hat{A}_1$ of the vector operator $\hat{A}$; they are defined in terms of the Cartesian coordinates $\hat{A}_x, \hat{A}_y, \hat{A}_z$ as follows:

$$\hat{A}_{\pm 1} = \mp\frac{1}{\sqrt{2}}(\hat{A}_x \pm i\hat{A}_y), \qquad \hat{A}_0 = \hat{A}_z. \tag{7.304}$$

For the particular case where $\hat{A}$ is equal to the position vector $\hat{R}$, we can express the components $\hat{R}_q$ (where $q = -1, 0, 1$),

$$\hat{R}_{\pm 1} = \mp\frac{1}{\sqrt{2}}(\hat{x} \pm i\hat{y}), \qquad \hat{R}_0 = \hat{z}, \tag{7.305}$$

in terms of the spherical coordinates (recall that $\hat{R}_1 = \hat{x} = r\sin\theta\cos\phi$, $\hat{R}_2 = \hat{y} = r\sin\theta\sin\phi$, and $\hat{R}_3 = \hat{z} = r\cos\theta$) as follows:

$$\hat{R}_{\pm 1} = \mp\frac{1}{\sqrt{2}} re^{\pm i\phi}\sin\theta, \qquad \hat{R}_0 = r\cos\theta. \tag{7.306}$$

Using the relations (7.291) to (7.293) and (7.301) to (7.303), we can ascertain that

$$\boxed{\left[\hat{J}_z, \hat{A}_q\right] = \hbar q\hat{A}_q \qquad (q = -1, 0, 1),} \tag{7.307}$$

$$\boxed{\left[\hat{J}_{\pm}, \hat{A}_q\right] = \hbar\sqrt{2 - q(q \pm 1)}\hat{A}_{q\pm 1} \qquad (q = -1, 0, 1).} \tag{7.308}$$

7.4.3 Transformation of Tensor Operators Under Rotations

In general, a tensor of rank k has 3^k components, where 3 denotes the dimension of the space. For instance, a tensor such as

$$\hat{T}_{ij} = A_i B_j \qquad (i, j = x, y, z), \tag{7.309}$$

which is equal to the product of the components of two vectors $\vec{A}$ and $\vec{B}$, is a *second-rank* tensor; this tensor has 3^2 components.

7.4.3.1 Transformation of Reducible Tensor Operators Under Rotations

A Cartesian tensor $\hat{T}_{ij}$ can be decomposed into three parts:

$$\hat{T}_{ij} = \hat{T}^{(0)}_{ij} + \hat{T}^{(1)}_{ij} + \hat{T}^{(2)}_{ij}, \tag{7.310}$$

with

$$\hat{T}^{(0)}_{ij} = \frac{1}{3}\delta_{i,j}\sum_{i=1}^{3}\hat{T}_{ii}, \tag{7.311}$$

$$\hat{T}^{(1)}_{ij} = \frac{1}{2}(\hat{T}_{ij} - \hat{T}_{ji}) \qquad (i \neq j), \tag{7.312}$$

$$\hat{T}^{(2)}_{ij} = \frac{1}{2}(\hat{T}_{ij} + \hat{T}_{ji}) - \hat{T}^{(0)}_{ij}. \tag{7.313}$$

Notice that if we add equations (7.311), (7.312), and (7.313), we end up with an identity relation: $\hat{T}_{ij} = \hat{T}_{ij}$.

The term $\hat{T}^{(0)}_{ij}$ has only one component and transforms like a scalar under rotations. The second term $\hat{T}^{(1)}_{ij}$ is an antisymmetric tensor of rank 1 which has three independent components; it transforms like a vector. The third term $\hat{T}^{(2)}_{ij}$ is a symmetric second-rank tensor with zero trace, and hence has five independent components; $\hat{T}^{(2)}_{ij}$ cannot be reduced further to tensors of lower rank. These five components define an irreducible second-rank tensor.

In general, any tensor of rank k can be decomposed into tensors of lower rank that are expressed in terms of linear combinations of its 3^k components. However, there always remain $(2k+1)$ components that behave as a tensor of rank k which cannot be reduced further. These $(2k+1)$ components are symmetric and traceless with respect to any two indices; they form the components of an *irreducible tensor* of rank k.

Equations (7.310) to (7.313) show how to decompose a Cartesian tensor operator, $\hat{T}_{ij}$, into a sum of irreducible *spherical* tensor operators $\hat{T}^{(0)}_{ij}, T^{(1)}_{ij}, T^{(2)}_{ij}$. Cartesian tensors are not very suitable for studying transformations under rotations, because they are reducible whenever their rank exceeds 1. In problems that display spherical symmetry, such as those encountered in subatomic physics, spherical tensors are very useful simplifying tools. It is therefore interesting to consider irreducible spherical tensor operators.

7.4.3.2 Transformation of Irreducible Tensor Operators Under Rotations

Let us now focus only on the representation of irreducible tensor operators in spherical coordinates. An irreducible spherical tensor operator of rank k (k is integer) is a set of $(2k+1)$ operators $T^{(k)}_q$, with $q = -k, \ldots, k$, which transform in the same way as angular momentum under a rotation of axes. For example, the case $k = 1$ corresponds to a vector. The quantities $T^{(1)}_q$ are related to the components of the vector $\hat{A}$ as follows (see (7.304)):

$$\hat{T}^{(1)}_{\pm 1} = \mp\frac{1}{\sqrt{2}}(\hat{A}_x \pm i\hat{A}_y), \qquad \hat{T}^{(1)}_0 = \hat{A}_z. \tag{7.314}$$

In what follows we are going to study some properties of spherical tensor operators and then determine how they transform under rotations.

First, let us look at the various commutation relations of spherical tensors with the angular momentum operator. Since a vector operator is a tensor of rank 1, we can rewrite equations (7.307) to (7.308), respectively, as follows:

$$\left[\hat{J}_z,\ \hat{T}_q^{(1)}\right] = \hbar q \hat{T}_q^{(1)} \qquad (q = -1, 0, 1), \tag{7.315}$$

$$\left[\hat{J}_\pm,\ \hat{T}_q^{(1)}\right] = \hbar\sqrt{1(1+1) - q(q\pm 1)}\,\hat{T}_{q\pm 1}^{(1)}, \tag{7.316}$$

where we have adopted the notation $\hat{A}_q = \hat{T}_q^{(1)}$. We can easily generalize these two relations to any spherical tensor of rank k, $\hat{T}_q^{(k)}$, and obtain these commutators:

$$\boxed{\left[\hat{J}_z,\ \hat{T}_q^{(k)}\right] = \hbar q \hat{T}_q^{(k)} \qquad (q = -k, -k+1, \ldots, k-1, k),} \tag{7.317}$$

$$\boxed{\left[\hat{J}_\pm,\ \hat{T}_q^{(k)}\right] = \hbar\sqrt{k(k+1) - q(q\pm 1)}\,\hat{T}_{q\pm 1}^{(k)}.} \tag{7.318}$$

Using the relations

$$\langle k,\ q' \mid \hat{J}_z \mid k,\ q\rangle = \hbar q \langle k,\ q' \mid k,\ q\rangle = \hbar q \delta_{q',q}, \tag{7.319}$$

$$\langle k,\ q' \mid \hat{J}_\pm \mid k,\ q\rangle = \hbar\sqrt{k(k+1) - q(q\pm 1)}\,\delta_{q',q\pm 1}, \tag{7.320}$$

along with (7.317) and (7.318), we can write

$$\sum_{q'=-k}^{k} \hat{T}_{q'}^{(k)} \langle k,\ q' \mid \hat{J}_z \mid k,\ q\rangle = \hbar q \hat{T}_q^{(k)} = \left[\hat{J}_z,\ \hat{T}_q^{(k)}\right], \tag{7.321}$$

$$\sum_{q'=-k}^{k} \hat{T}_{q'}^{(k)} \langle k,\ q' \mid \hat{J}_\pm \mid k,\ q\rangle = \hbar\sqrt{k(k+1) - q(q\pm 1)}\,\hat{T}_{q\pm 1}^{(k)} = \left[\hat{J}_\pm,\ \hat{T}_q^{(k)}\right]. \tag{7.322}$$

The previous two relations can be combined into

$$\boxed{\left[\hat{\vec{J}},\ \hat{T}_q^{(k)}\right] = \sum_{q'=-k}^{k} \hat{T}_{q'}^{(k)} \langle k,\ q' \mid \hat{\vec{J}} \mid k,\ q\rangle} \tag{7.323}$$

or

$$\boxed{\left[\vec{n}\cdot\hat{\vec{J}},\ \hat{T}_q^{(k)}\right] = \sum_{q'=-k}^{k} \hat{T}_{q'}^{(k)} \langle k,\ q' \mid \vec{n}\cdot\hat{\vec{J}} \mid k,\ q\rangle.} \tag{7.324}$$

Having determined the commutation relations of the tensor operators with the angular momentum (7.324), we are now well equipped to study how irreducible spherical tensor operators transform under rotations. Using (7.285), we can write the transformation relation of a spherical tensor $T_q^{(k)}$ under an infinitesimal rotation as follows:

$$\hat{R}_{\vec{n}}(\delta\theta)\,\hat{T}_q^{(k)}\,\hat{R}_{\vec{n}}^{\dagger}(\delta\theta) = T_q^{(k)} - \frac{i}{\hbar}\,\delta\theta\left[\vec{n}\cdot\hat{\vec{J}},\ \hat{T}_q^{(k)}\right]. \tag{7.325}$$

Inserting (7.324) into (7.325), we obtain

$$\hat{R}(\delta\theta)\,\hat{T}^{(k)}_q\,\hat{R}^\dagger(\delta\theta) = \sum_{q'=-k}^{k} \hat{T}^{(k)}_{q'} \langle k,\, q' \mid 1 - \frac{i}{\hbar}\,\delta\theta\,\vec{n}\cdot\hat{\vec{J}} \mid k,\, q\rangle, \tag{7.326}$$

which can be reduced to

$$\boxed{\hat{R}(\delta\theta)\,\hat{T}^{(k)}_q\,\hat{R}^\dagger(\delta\theta) = \sum_{q'} \hat{T}^{(k)}_{q'} \langle k,\, q' \mid e^{-i\delta\theta\,\vec{n}\cdot\hat{\vec{J}}/\hbar} \mid k,\, q\rangle.} \tag{7.327}$$

This result also holds for finite (Euler) rotations

$$\boxed{\hat{R}(\alpha,\beta,\gamma)\,\hat{T}^{(k)}_q\,\hat{R}^\dagger(\alpha,\beta,\gamma) = \sum_{q'=-k}^{k} \hat{T}^{(k)}_{q'} \langle k,\, q' \mid \hat{R}(\alpha,\beta,\gamma) \mid k,\, q\rangle = \sum_{q'} \hat{T}^{(k)}_{q'}\, D^{(k)}_{q'q}(\alpha,\beta,\gamma).} \tag{7.328}$$

7.4.4 Wigner–Eckart Theorem for Spherical Tensor Operators

Taking the matrix elements of (7.317) between eigenstates of $\hat{\vec{J}}^2$ and $\hat{J}_z$, we find

$$\langle j',\, m' \mid \left[\hat{J}_z,\, \hat{T}^{(k)}_q\right] - \hbar q \hat{T}^{(k)}_q \mid j,\, m\rangle = 0 \tag{7.329}$$

or

$$(m' - m - q)\langle j',\, m' \mid \hat{T}^{(k)}_q \mid j,\, m\rangle = 0. \tag{7.330}$$

This implies that $\langle j',\, m' \mid \hat{T}^{(k)}_q \mid j,\, m\rangle$ vanishes unless $m' = m + q$. This property suggests that the quantity $\langle j',\, m' \mid \hat{T}^{(k)}_q \mid j,\, m\rangle$ must be proportional to the Clebsch–Gordan coefficient $\langle j',\, m' \mid j,k;\, m,q\rangle$; hence (7.330) leads to

$$(m' - m - q)\langle j',\, m' \mid j,k;\, m,q\rangle = 0. \tag{7.331}$$

Now, taking the matrix elements of (7.318) between $\mid j,\, m\rangle$ and $\mid j',\, m'\rangle$, we obtain

$$\begin{aligned}
&\sqrt{(j' \pm m')(j' \mp m' + 1)}\,\langle j', m' \mp 1 | \hat{T}^{(k)}_q \mid j,\, m\rangle \\
&\quad = \sqrt{(j \mp m)(j \pm m + 1)}\,\langle j', m' | \hat{T}^{(k)}_q \mid j,\, m \pm 1\rangle \\
&\quad + \sqrt{(k \mp q)(k \pm q + 1)}\,\langle j', m' | \hat{T}^{(k)}_{q\pm1} \mid j,\, m\rangle.
\end{aligned} \tag{7.332}$$

This equation has a structure which is identical to the recursion relation (7.158). For instance, substituting $j = j', m = m', j_1 = j, m_1 = m, j_2 = k, m_2 = q$ into (7.158), we end up with

$$\begin{aligned}
&\sqrt{(j' \pm m')(j' \mp m' + 1)}\,\langle j', m' \mp 1 \mid j,k;\, m,q\rangle \\
&\quad = \sqrt{(j \mp m)(j \pm m + 1)}\,\langle j', m' \mid j,k;\, m \pm 1, q\rangle \\
&\quad + \sqrt{(k \mp q)(k \pm q + 1)}\,\langle j', m' \mid j,k;\, m, q \pm 1\rangle.
\end{aligned} \tag{7.333}$$

A comparison of (7.330) with (7.331) and (7.332) with (7.333) suggests that the dependence of $\langle j', m'|T_q^{(k)} \mid j,\ m\rangle$ on m', m, q is through a Clebsch–Gordan coefficient. The dependence, however, of $\langle j', m'|T_q^{(k)} \mid j,\ m\rangle$ on j', j, k has yet to be determined.

We can now state the **Wigner–Eckart theorem:** *The matrix elements of spherical tensor operators* $\hat{T}_q^{(k)}$ *with respect to angular momentum eigenstates* $\mid j,\ m\rangle$ are given by

$$\boxed{\langle j', m'|\hat{T}_q^{(k)} \mid j,\ m\rangle = \langle j, k;\ m, q|j',\ m'\rangle\langle j' \parallel \hat{T}^{(k)} \parallel j\rangle.} \tag{7.334}$$

The factor $\langle j' \parallel \hat{T}^{(k)} \parallel j\rangle$, which depends only on j', j, k, is called the *reduced matrix element* of the tensor $\hat{T}_q^{(k)}$ (note that the double bars notation is used to distinguish the reduced matrix elements, $\langle j' \parallel \hat{T}^{(k)} \parallel j\rangle$, from the matrix elements, $\langle j',\ m'|\hat{T}_q^{(k)} \mid j,\ m\rangle$). The theorem implies that the matrix elements $\langle j',\ m'|\hat{T}_q^{(k)} \mid j,\ m\rangle$ are written as the product of two terms: a Clebsch–Gordan coefficient $\langle j, k;\ m, q|j',\ q'\rangle$—which depends on the geometry of the system (i.e., the orientation of the system with respect to the z-axis), but not on its dynamics (i.e., j', j, k)—and a dynamical factor, the reduced matrix element, which does not depend on the orientation of the system in space (m', q, m). The quantum numbers m', m, q—which specify the projections of the angular momenta $\hat{\vec{J}}'$, $\hat{\vec{J}}$, and $\vec{k}$ onto the z-axis—give the orientation of the system in space, for they specify its orientation with respect to the z-axis. As for j', j, k, they are related to the dynamics of the system, not to its orientation in space.

Wigner–Eckart theorem for a scalar operator

The simplest application of the Wigner–Eckart theorem is when dealing with a scalar operator $\hat{B}$. As seen above, a scalar is a tensor of rank $k = 0$; hence $q = 0$ as well; thus, equation (7.334) yields

$$\langle j', m'|\hat{B} \mid j,\ m\rangle = \langle j, 0;\ m, 0|j',\ m'\rangle\langle j' \parallel \hat{B} \parallel j\rangle = \langle j' \parallel \hat{B} \parallel j\rangle\delta j' j\,\delta m' m, \tag{7.335}$$

since $\langle j, 0;\ m, 0|j',\ m'\rangle = \delta j' j\,\delta m' m$.

Wigner–Eckart theorem for a vector operator

As shown in (7.314), a vector is a tensor of rank 1: $T^{(1)} = A^{(1)} = \hat{\vec{A}}$, with $A_0^{(1)} = A_0 = A_z$ and $A_{\pm 1}^{(1)} = A_{\pm 1} = \mp(\hat{A}_x \pm i\hat{A}_y)/\sqrt{2}$. An application of (7.334) to the q-component of a vector operator $\hat{\vec{A}}$ leads to

$$\langle j', m'|\hat{A}_q \mid j,\ m\rangle = \langle j, 1;\ m, q|j',\ m'\rangle\langle j' \parallel \hat{\vec{A}} \parallel j\rangle. \tag{7.336}$$

For instance, in the case of the angular momentum $\hat{\vec{J}}$, we have

$$\langle j', m'|\hat{J}_q \mid j,\ m\rangle = \langle j, 1;\ m, q|j',\ m'\rangle\langle j' \parallel \hat{\vec{J}} \parallel j\rangle. \tag{7.337}$$

Applying this relation to the component $\hat{J}_0$,

$$\langle j', m'|\hat{J}_0 \mid j,\ m\rangle = \langle j, 1;\ m, 0|j',\ m'\rangle\langle j' \parallel \hat{\vec{J}} \parallel j\rangle. \tag{7.338}$$

Since $\langle j', m'|\hat{J}_0 \mid j,\ m\rangle = \hbar m\,\delta j' j\,\delta m' m$ and the coefficient $\langle j, 1;\ m, 0 \mid j,\ m\rangle$ is equal to $\langle j, 1;\ m, 0 \mid j,\ m\rangle = m/\sqrt{j(j+1)}$, we have

$$\hbar m\,\delta j' j\,\delta m' m = \frac{m}{\sqrt{j(j+1)}}\langle j' \parallel \hat{\vec{J}} \parallel j\rangle \quad\Longrightarrow\quad \langle j' \parallel \hat{\vec{J}} \parallel j\rangle = \hbar\sqrt{j(j+1)}\,\delta_{j',j}. \tag{7.339}$$

Due to the selection rules imposed by the Clebsch–Gordan coefficients, we see from (7.336) that *a spin zero particle cannot have a dipole moment*. Since $\langle 0,1;\ 0,q|0,\ 0\rangle = 0$, we have $\langle 0,\ 0|\hat{L}_q \mid 0,\ 0\rangle = \langle 0,1;\ 0,q|0,\ 0\rangle\langle 0 \parallel \hat{\vec{L}} \parallel 0\rangle = 0$; the dipole moment is $\hat{\vec{\mu}} = -q\hat{\vec{L}}/(2mc)$. Similarly, *a spin $\frac{1}{2}$ particle cannot have a quadrupole moment*, because as $\langle \frac{1}{2},2;\ m,q|\frac{1}{2},\ m'\rangle = 0$, we have $\langle \frac{1}{2},\ m'|\hat{T}^{(2)}_q|\frac{1}{2},\ m\rangle = \langle \frac{1}{2},2;\ m,q|\frac{1}{2},\ m'\rangle\langle \frac{1}{2} \parallel \hat{T}^{(2)} \parallel \frac{1}{2}\rangle = 0$.

Wigner–Eckart theorem for a scalar product $\hat{\vec{J}}\cdot\hat{\vec{A}}$

On the one hand, since $\hat{\vec{J}}\cdot\hat{\vec{A}} = \hat{J}_0\hat{A}_0 - \hat{J}_{+1}\hat{A}_{-1} - \hat{J}_{-1}\hat{A}_{+1}$ and since $\hat{J}_0 \mid j,\ m\rangle = \hbar m \mid j,\ m\rangle$ and $\hat{J}_{\pm 1} \mid j,\ m\rangle = \mp(\hbar/\sqrt{2})\sqrt{j(j+1) - m(m\pm 1)}|j,\ m\pm 1\rangle$, we have

$$\langle j,m|\hat{\vec{J}}\cdot\hat{\vec{A}} \mid j,m\rangle = \hbar m\langle j,m \mid \hat{A}_0 \mid j,m\rangle - \frac{\hbar}{\sqrt{2}}\sqrt{j(j+1) - m(m+1)}\langle j,m+1|\hat{A}_{+1} \mid j,m\rangle + \frac{\hbar}{\sqrt{2}}\sqrt{j(j+1) - m(m-1)}\langle j,m-1|\hat{A}_{+1} \mid j,m\rangle. \tag{7.340}$$

On the other hand, from the Wigner–Eckart theorem (7.334) we have $\langle j,\ m \mid \hat{A}_0 \mid j,\ m\rangle = \langle j,1;\ m,0 \mid j,\ m\rangle\langle j \parallel \hat{\vec{A}} \parallel j\rangle$, $\langle j,\ m+1|\hat{A}_{+1} \mid j,\ m\rangle = \langle j,1;\ m,1|j,\ m+1\rangle\langle j \parallel \hat{\vec{A}} \parallel j\rangle$ and $\langle j,m-1|\hat{A}_{-1} \mid j,\ m\rangle = \langle j,1;\ m,-1|j,\ m-1\rangle\langle j \parallel \hat{\vec{A}} \parallel j\rangle$; substituting these terms into (7.340) we obtain

$$\langle j,\ m|\hat{\vec{J}}\cdot\hat{\vec{A}} \mid j,\ m\rangle = \Big[\hbar m\langle j,1;\ m,0 \mid j,\ m\rangle - \frac{\hbar}{\sqrt{2}}\langle j,1;\ m,1|j,\ m+1\rangle\sqrt{j(j+1) - m(m+1)} + \frac{\hbar}{\sqrt{2}}\langle j,1;\ m,-1|j,\ m-1\rangle\sqrt{j(j+1) - m(m-1)}\Big]\langle j \parallel \hat{\vec{A}} \parallel j\rangle. \tag{7.341}$$

When $\hat{\vec{A}} = \hat{\vec{J}}$ this relation leads to

$$\langle j,\ m|\hat{\vec{J}}^2 \mid j,\ m\rangle = \Big[\hbar m\langle j,1;\ m,0 \mid j,\ m\rangle - \frac{\hbar}{\sqrt{2}}\langle j,1;\ m,1|j,\ m+1\rangle\sqrt{j(j+1) - m(m+1)} + \frac{\hbar}{\sqrt{2}}\langle j,1;\ m,-1|j,\ m-1\rangle\sqrt{j(j+1) - m(m-1)}\Big]\langle j \parallel \hat{\vec{J}} \parallel j\rangle. \tag{7.342}$$

We are now equipped to obtain a relation between the matrix elements of a vector operator $\hat{\vec{A}}$ and the matrix elements of the scalar operator $\hat{\vec{J}}\cdot\hat{\vec{A}}$; this relation is useful in the calculation of the hydrogen's energy corrections due to the Zeeman effect (see Chapter 9). For this, we need to calculate two ratios: the first is between (7.336) and (7.337)

$$\frac{\langle j,\ m'|\hat{A}_q \mid j,\ m\rangle}{\langle j,\ m'|\hat{J}_q \mid j,\ m\rangle} = \frac{\langle j \parallel \hat{\vec{A}} \parallel j\rangle}{\langle j \parallel \hat{\vec{J}} \parallel j\rangle} \tag{7.343}$$

and the second is between (7.341) and (7.342)

$$\frac{\langle j,\, m|\hat{\vec{J}}\cdot\hat{\vec{A}}\,|\,j,\, m\rangle}{\langle j,\, m|\hat{\vec{J}}^2\,|\,j,\, m\rangle} = \frac{\langle j\,\|\,\hat{\vec{A}}\,\|\,j\rangle}{\langle j\,\|\,\hat{\vec{J}}\,\|\,j\rangle} \quad\Longrightarrow\quad \frac{\langle j\, m\,|\,\hat{\vec{J}}\cdot\hat{\vec{A}}\,|\,j,\, m\rangle}{\hbar^2 j(j+1)} = \frac{\langle j\,\|\,\hat{\vec{A}}\,\|\,j\rangle}{\langle j\,\|\,\hat{\vec{J}}\,\|\,j\rangle}, \tag{7.344}$$

since $\langle j\, m\,|\,\hat{\vec{J}}^2\,|\,j,\, m\rangle = \hbar^2 j(j+1)$. Equating (7.343) and (7.344) we obtain

$$\boxed{\langle j,\, m'|\hat{A}_q\,|\,j,\, m\rangle = \frac{\langle j,\, m|\hat{\vec{J}}\cdot\hat{\vec{A}}\,|\,j,\, m\rangle}{\hbar^2 j(j+1)}\langle j,\, m'|\hat{J}_q\,|\,j,\, m\rangle.} \tag{7.345}$$

An important application of this relation pertains to the case where the vector operator $\hat{\vec{A}}$ is a spin angular momentum $\hat{\vec{S}}$. Since

$$\begin{aligned}\hat{\vec{J}}\cdot\hat{\vec{S}} &= (\hat{\vec{L}}+\hat{\vec{S}})\cdot\hat{\vec{S}} = \hat{\vec{L}}\cdot\hat{\vec{S}}+\hat{S}^2 = \frac{(\hat{\vec{L}}+\hat{\vec{S}})^2-\hat{L}^2-\hat{S}^2}{2}+\hat{S}^2 = \frac{\hat{J}^2-\hat{L}^2-\hat{S}^2}{2}+\hat{S}^2 \\ &= \frac{\hat{J}^2-\hat{L}^2+\hat{S}^2}{2},\end{aligned} \tag{7.346}$$

and since $|\,j,\, m\rangle$ is a joint eigenstate of $\hat{J}^2$, $\hat{L}^2$, $\hat{S}^2$ and $\hat{J}_z$ with eigenvalues $\hbar^2 j(j+1)$, $\hbar^2 l(l+1)$, $\hbar^2 s(s+1)$, and $\hbar m$, respectively, the matrix element of $\hat{S}_z$ then becomes easy to calculate from (7.345):

$$\langle j, m|\hat{S}_z\,|\,j, m\rangle = \frac{\langle j, m\,|\,\hat{\vec{J}}\cdot\hat{\vec{S}}\,|\,j, m\rangle}{\hbar^2 j(j+1)}\langle j, m|\hat{J}_z\,|\,j, m\rangle = \frac{j(j+1)-l(l+1)+s(s+1)}{2j(j+1)}\hbar m. \tag{7.347}$$

7.5 Solved Problems

Problem 7.1

(a) Consider the operator $\hat{R}_z = e^{i\alpha\hat{J}_z/\hbar}$, representing a rotation of (finite) angle α about the z-axis. Show how $\hat{J}_x$ and $\hat{J}_y$ transform under $\hat{R}_z$. Using these results, determine how the angular momentum operator $\hat{\vec{J}}$ transform under the rotation.

(b) Show how a vector operator $\hat{\vec{A}}$ transforms under a rotation of angle α about the y-axis.

(c) Show that $e^{i\pi\hat{J}_z/\hbar}e^{i\alpha\hat{J}_y/\hbar}e^{-i\pi\hat{J}_z/\hbar} = e^{-i\alpha\hat{J}_y/\hbar}$.

Solution

(a) Under this rotation, an operator $\hat{B}$ transforms like $\hat{B}' = \hat{R}_z\hat{B}\hat{R}_z^\dagger = e^{i\alpha\hat{J}_z/\hbar}\hat{B}e^{-i\alpha\hat{J}_z/\hbar}$ (see (7.23)). Using the relation

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + \big[\hat{A},\hat{B}\big] + \frac{1}{2!}\big[\hat{A},\big[\hat{A},\hat{B}\big]\big] + \frac{1}{3!}\big[\hat{A},\big[\hat{A},\big[\hat{A},\hat{B}\big]\big]\big] + \cdots, \tag{7.348}$$

along with the commutation relations $\left[\hat{J}_z, \hat{J}_y\right] = -i\hbar\hat{J}_x$ and $\left[\hat{J}_z, \hat{J}_x\right] = i\hbar\hat{J}_y$, we have

$$\begin{aligned} e^{i\alpha\hat{J}_z/\hbar}\hat{J}_x e^{-i\alpha\hat{J}_z/\hbar} &= \hat{J}_x + \frac{i\alpha}{\hbar}\left[\hat{J}_z, \hat{J}_x\right] - \frac{\alpha^2}{2!\hbar^2}\left[\hat{J}_z, \left[\hat{J}_z, \hat{J}_x\right]\right] \\ &\quad - \frac{i\alpha^3}{3!\hbar^3}\left[\hat{J}_z, \left[\hat{J}_z, \left[\hat{J}_z, \hat{J}_x\right]\right]\right] + \cdots \\ &= \hat{J}_x - \alpha\hat{J}_y - \frac{\alpha^2}{2!}\hat{J}_x + \frac{\alpha^3}{3!}\hat{J}_y + \frac{\alpha^4}{4!}\hat{J}_x - \frac{\alpha^5}{5!}\hat{J}_y + \cdots \\ &= \hat{J}_x\left(1 - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} + \cdots\right) - \hat{J}_y\left(\alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \cdots\right) \\ &= \hat{J}_x \cos\alpha - \hat{J}_y \sin\alpha. \end{aligned} \tag{7.349}$$

Similarly, we can show that

$$e^{i\alpha\hat{J}_z/\hbar}\hat{J}_y e^{-i\alpha\hat{J}_z/\hbar} = \hat{J}_y \cos\alpha + \hat{J}_x \sin\alpha. \tag{7.350}$$

As $\hat{J}_z$ is invariant under an arbitrary rotation about the z-axis ($e^{i\alpha\hat{J}_z/\hbar}\hat{J}_z e^{-i\alpha\hat{J}_z/\hbar} = \hat{J}_z$), we can condense equations (7.349) and (7.350) into a single matrix relation:

$$e^{i\alpha\hat{J}_z/\hbar}\hat{\vec{J}}e^{-i\alpha\hat{J}_z/\hbar} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \hat{J}_x \\ \hat{J}_y \\ \hat{J}_z \end{pmatrix}. \tag{7.351}$$

(b) Using the commutation relations $\left[\hat{J}_y, \hat{A}_x\right] = -i\hbar\hat{A}_z$ and $\left[\hat{J}_y, \hat{A}_z\right] = i\hbar\hat{A}_x$ (see (7.291) to (7.293)) along with (7.348), we have

$$\begin{aligned} e^{i\alpha\hat{J}_y/\hbar}\hat{A}_x e^{-i\alpha\hat{J}_y/\hbar} &= \hat{A}_x + \frac{i\alpha}{\hbar}\left[\hat{J}_y, \hat{A}_x\right] - \frac{\alpha^2}{2!\hbar^2}\left[\hat{J}_y, \left[\hat{J}_y, \hat{A}_x\right]\right] \\ &\quad - \frac{i\alpha^3}{3!\hbar^3}\left[\hat{J}_y, \left[\hat{J}_y, \left[\hat{J}_y, \hat{A}_x\right]\right]\right] + \cdots \\ &= \hat{A}_x + \alpha\hat{A}_z - \frac{\alpha^2}{2!}\hat{A}_x - \frac{\alpha^3}{3!}\hat{A}_z + \frac{\alpha^4}{4!}\hat{A}_x + \frac{\alpha^5}{5!}\hat{J}_z + \cdots \\ &= \hat{A}_x\left(1 - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} + \cdots\right) + \hat{A}_z\left(\alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} + \cdots\right) \\ &= \hat{A}_x \cos\alpha + \hat{A}_z \sin\alpha. \end{aligned} \tag{7.352}$$

Similarly, we can show that

$$\hat{A}_z' = e^{i\alpha\hat{J}_y/\hbar}\hat{A}_z e^{-i\alpha\hat{J}_y/\hbar} = -\hat{A}_x \sin\alpha + \hat{A}_z \cos\alpha. \tag{7.353}$$

Also, since $\hat{A}_y$ is invariant under an arbitrary rotation about the y-axis, we may combine equations (7.352) and (7.353) to find the vector operator $\hat{\vec{A}}'$ obtained by rotating $\hat{\vec{A}}$ through an angle α about the y-axis:

$$\hat{\vec{A}}' = e^{i\alpha\hat{J}_y/\hbar}\hat{\vec{A}}e^{-i\alpha\hat{J}_y/\hbar} = \begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix}\begin{pmatrix} \hat{A}_x \\ \hat{A}_y \\ \hat{A}_z \end{pmatrix}. \tag{7.354}$$

(c) Expanding $e^{i\alpha \hat{J}_y/\hbar}$ and then using (7.350), we obtain

$$\begin{aligned} e^{i\pi \hat{J}_z/\hbar} e^{i\alpha \hat{J}_y/\hbar} e^{-i\pi \hat{J}_z/\hbar} &= \sum_{n=0}^{\infty} \frac{(i\alpha/\hbar)^n}{n!} e^{i\pi \hat{J}_z/\hbar} \left(\hat{J}_y\right)^n e^{-i\pi \hat{J}_z/\hbar} \\ &= \sum_{n=0}^{\infty} \frac{(i\alpha/\hbar)^n}{n!} \left(\hat{J}_y \cos\pi + \hat{J}_x \sin\pi\right)^n = \sum_{n=0}^{\infty} \frac{(-i\alpha/\hbar)^n}{n!} \left(\hat{J}_y\right)^n \\ &= e^{-i\alpha \hat{J}_y/\hbar}. \end{aligned} \tag{7.355}$$

Problem 7.2

Use the Pauli matrices $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, and $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, to show that

(a) $e^{-i\alpha\sigma_x} = I\cos\alpha - i\sigma_x \sin\alpha$, where I is the unit matrix,

(b) $e^{i\alpha\sigma_x}\sigma_z e^{-i\alpha\sigma_x} = \sigma_z \cos(2\alpha) + \sigma_y \sin(2\alpha)$.

Solution

(a) Using the expansion

$$e^{-i\alpha\sigma_x} = \sum_{n=0}^{\infty} \frac{(-i)^{2n}}{(2n)!} (\alpha)^{2n} \sigma_x^{2n} + \sum_{n=0}^{\infty} \frac{(-i)^{2n+1}}{(2n+1)!} (\alpha)^{2n+1} \sigma_x^{2n+1}, \tag{7.356}$$

and since $\sigma_x^2 = I$, $\sigma_x^{2n} = I$, and $\sigma_x^{2n+1} = \sigma_x$, where I is the unit matrix, we have

$$\begin{aligned} e^{-i\alpha\sigma_x} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} (\alpha)^{2n} - i\sigma_x \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} (\alpha)^{2n+1} \\ &= I\cos\alpha - i\sigma_x \sin\alpha. \end{aligned} \tag{7.357}$$

(b) From (7.357) we can write

$$\begin{aligned} e^{i\alpha\sigma_x}\sigma_z e^{-i\alpha\sigma_x} &= (\cos\alpha + i\sigma_x \sin\alpha)\sigma_z(\cos\alpha - i\sigma_x \sin\alpha) \\ &= \sigma_z \cos^2\alpha + \sigma_x\sigma_z\sigma_x \sin^2\alpha + i[\sigma_x, \sigma_z] \sin\alpha\cos\alpha, \end{aligned} \tag{7.358}$$

which, when using the facts that $\sigma_x\sigma_z = -\sigma_z\sigma_x$, $\sigma_x^2 = I$, and $[\sigma_x, \sigma_z] = -2i\sigma_y$, reduces to

$$\begin{aligned} e^{i\alpha\sigma_x}\sigma_z e^{-i\alpha\sigma_x} &= \sigma_z \cos^2\alpha - \sigma_z\sigma_x^2 \sin^2\alpha + 2\sigma_y \sin\alpha\cos\alpha \\ &= \sigma_z(\cos^2\alpha - \sin^2\alpha) + \sigma_y \sin(2\alpha) \\ &= \sigma_z \cos(2\alpha) + \sigma_y \sin(2\alpha). \end{aligned} \tag{7.359}$$

Problem 7.3

Find the Clebsch–Gordan coefficients associated with the addition of two angular momenta $j_1 = 1$ and $j_2 = 1$.

Solution

The addition of $j_1 = 1$ and $j_2 = 1$ is encountered, for example, in a two-particle system where the angular momenta of both particles are orbital.

The allowed values of the total angular momentum are between $|j_1 - j_2| \leq j \leq j_1 + j_2$; hence $j = 0, 1, 2$. To calculate the relevant Clebsch–Gordan coefficients, we need to find the basis vectors $\{|j, m\rangle\}$, which are common eigenvectors of $\hat{\vec{J}}_1^2$, $\hat{\vec{J}}_2^2$, $\hat{\vec{J}}^2$ and $\hat{J}_z$, in terms of $\{|1,1;\, m_1, m_2\rangle\}$.

Eigenvectors $|j, m\rangle$ associated with $j = 2$

The state $|2, 2\rangle$ is simply given by

$$\boxed{|2, 2\rangle = |1,1;\, 1,1\rangle;} \tag{7.360}$$

the corresponding Clebsch–Gordan coefficient is thus given by $\langle 1,1;\, 1,1 \,|\, 2,\, 2\rangle = 1$.

As for $|2, 1\rangle$, it can be found by applying J_- to $|2, 2\rangle$ and $(J_{1-} + J_{2-})$ to $|1,1;\, 1,1\rangle$, and then equating the two results

$$J_- |2, 2\rangle = (J_{1-} + J_{2-})|1,1;\, 1,1\rangle. \tag{7.361}$$

This leads to

$$2\hbar\, |2, 1\rangle = \sqrt{2}\hbar\Big(|1,1;\, 1,0\rangle + |1,1;\, 0,1\rangle\Big) \tag{7.362}$$

or to

$$\boxed{|2, 1\rangle = \frac{1}{\sqrt{2}}\Big(|1,1;\, 1,0\rangle + |1,1;\, 0,1\rangle\Big);} \tag{7.363}$$

hence $\langle 1,1;\ 1,0 \,|\, 2,\, 1\rangle = \langle 1,1;\, 0,1 \,|\, 2,\, 1\rangle = 1/\sqrt{2}$. Using (7.363), we can find $|2,\, 0\rangle$ by applying J_- to $|2,\, 1\rangle$ and $(J_{1-} + J_{2-})$ to $[|1,1;\, 1,0\rangle + |1,1;\, 0,1\rangle]$:

$$J_- |2, 1\rangle = \frac{1}{\sqrt{2}}\hbar\,(J_{1-} + J_{2-})\,[|1,1;\, 1,0\rangle + |1,1;\, 0,1\rangle]\,, \tag{7.364}$$

which leads to

$$\boxed{|2, 0\rangle = \frac{1}{\sqrt{6}}\Big(|1,1;\, 1,-1\rangle + 2|1,1;\, 0,0\rangle + |1,1;\, -1,1\rangle\Big);} \tag{7.365}$$

hence $\langle 1,1;\, 1,-1 \,|\, 2,\, 0\rangle = \langle 1,1;\, -1,1 \,|\, 2,\, 0\rangle = 1/\sqrt{6}$ and $\langle 1,1;\, 0,0 \,|\, 2,\, 0\rangle = 2/\sqrt{6}$.

Similarly, by repeated applications of J_- and $(J_{1-} + J_{2-})$, we can show that

$$\boxed{|2, -1\rangle = \frac{1}{\sqrt{2}}\Big(|1,1;\, 0,-1\rangle + |1,1;\, -1,0\rangle\Big),} \tag{7.366}$$

$$\boxed{|2, -2\rangle = |1,1;\, -1,-1\rangle,} \tag{7.367}$$

with $\langle 1,1;\, 0,-1|2,\, -1\rangle = \langle 1,1;\, -1,0|2,\, -1\rangle = 1/\sqrt{2}$ and $\langle 1,1;\, -1,-1|2,\, -2\rangle = 1$.

Eigenvectors $|\, j, m\rangle$ associated with $j = 1$

The relation

$$|1, m\rangle = \sum_{m_1=-1}^{1} \sum_{m_2=-1}^{1} \langle 1,1;\, m_1, m_2|1,\, m\rangle |1,1;\, m_1, m_2\rangle \tag{7.368}$$

leads to

$$|\,1,\,1\rangle = a|1,1;\,1,0\rangle + b|1,1;\,0,1\rangle, \tag{7.369}$$

where $a = \langle 1,1;\,1,0\,|\,1,\,1\rangle$ and $b = \langle 1,1;\,0,1|1,\,1\rangle$. Since $|\,1,\,1\rangle$, $|1,1;\,1,0\rangle$ and $|1,1;\,0,1\rangle$ are all normalized, and since $|1,1;\,1,0\rangle$ is orthogonal to $|1,1;\,0,1\rangle$ and a and b are real, we have

$$\langle 1,1\,|\,1,\,1\rangle = a^2 + b^2 = 1. \tag{7.370}$$

Now, since $\langle 2,1\,|\,1,\,1\rangle = 0$, equations (7.363) and (7.369) yield

$$\langle 2,1\,|\,1,\,1\rangle = \frac{a}{\sqrt{2}} + \frac{b}{\sqrt{2}} = 0. \tag{7.371}$$

A combination of (7.370) and (7.371) leads to $a = -b = \pm 1/\sqrt{2}$. The signs of a and b have yet to be found. The phase convention mandates that coefficients like $\langle j_1, j_2;\, j_1, (j - j_1)|j,\, j\rangle$ must be positive. Thus, we have $a = 1/\sqrt{2}$ and $b = -1/\sqrt{2}$, which when inserted into (7.369) give

$$\boxed{|\,1,\,1\rangle = \frac{1}{\sqrt{2}}\Big(|1,1;\,1,0\rangle - |1,1;\,0,1\rangle\Big).} \tag{7.372}$$

This yields $\langle 1,1;\,1,0\,|\,1,\,1\rangle = \frac{1}{2}$ and $\langle 1,1;\,0,1\,|\,1,\,1\rangle = -\frac{1}{2}$.

To find $|1,\,0\rangle$ we proceed as we did above when we obtained the states $|\,2,\,1\rangle$, $|\,2,\,0\rangle$, ..., $|2,\,-2\rangle$ by repeatedly applying J_- on $|\,2,\,2\rangle$. In this way, the application of J_- on $|\,1,\,1\rangle$ and $(J_{1-} + J_{2-})$ on $[|1,1;\,1,0\rangle - |1,1;\,0,1\rangle]$,

$$J_-\,|\,1,\,1\rangle = \frac{1}{2}\,(J_{1-} + J_{2-})\,[|1,1;\,1,0\rangle - |1,1;\,0,1\rangle] \tag{7.373}$$

gives

$$\sqrt{2}\hbar\,|\,1,\,0\rangle = \frac{\sqrt{2}\hbar}{2}\,[|1,1;\,1,-1\rangle - |1,1;\,-1,1\rangle], \tag{7.374}$$

or

$$\boxed{|\,1,\,0\rangle = \frac{1}{\sqrt{2}}\Big(|1,1;\,1,-1\rangle - |1,1;\,-1,1\rangle\Big),} \tag{7.375}$$

with $\langle 1,1;\,1,-1\,|\,1,\,0\rangle = \frac{1}{\sqrt{2}}$ and $\langle 1,1;\,-1,1\,|\,1,\,0\rangle = -1/\sqrt{2}$.

Similarly, we can show that

$$\boxed{|\,1,\,-1\rangle = \frac{1}{\sqrt{2}}\Big(|1,1;\,0,-1\rangle - |1,1;\,-1,0\rangle\Big);} \tag{7.376}$$

hence $\langle 1,1;\,0,-1\,|\,1,\,-1\rangle = 1/\sqrt{2}$ and $\langle 1,1;\,-1,0\,|\,1,\,-1\rangle = -1/\sqrt{2}$.

Eigenvector $|0,\,0\rangle$ associated with $j = 0$

Since

$$|0,\,0\rangle = a|1,1;\,1,-1\rangle + b|1,1;\,0,0\rangle + c|1,1;\,-1,1\rangle, \tag{7.377}$$

where $a = \langle 1,1;\ 1,-1 \mid 0,\ 0\rangle$, $b = \langle 1,1;\ 0,0 \mid 0,\ 0\rangle$, and $c = \langle 1,1;\ -1,1 \mid 0,\ 0\rangle$ are real, and since the states $|0,\ 0\rangle$, $|1,1;\ 1,-1\rangle$, $|1,1;\ 0,0\rangle$, and $|1,1;\ -1,1\rangle$ are normal, we have

$$\langle 0,\ 0 \mid 0,\ 0\rangle = a^2 + b^2 + c^2 = 1. \tag{7.378}$$

Now, combining (7.365), (7.375), and (7.377), we obtain

$$\langle 2,\ 0 \mid 0,\ 0\rangle = \frac{a}{\sqrt{6}} + \frac{2b}{\sqrt{6}} + \frac{c}{\sqrt{6}} = 0, \tag{7.379}$$

$$\langle 1,\ 0 \mid 0,\ 0\rangle = \frac{a}{\sqrt{2}} - \frac{c}{\sqrt{2}} = 0. \tag{7.380}$$

Since a is by convention positive, we can show that the solutions of (7.378), (7.379), and (7.380) are given by $a = 1/\sqrt{3}$, $b = -1/\sqrt{3}$, $c = 1/\sqrt{3}$, and consequently

$$\boxed{|0,\ 0\rangle = \frac{1}{\sqrt{3}}\Big(|1,1;\ 1,-1\rangle - |1,1;\ 0,0\rangle + |1,1;\ -1,1\rangle\Big),} \tag{7.381}$$

with $\langle 1,1;\ 1,-1 \mid 0,\ 0\rangle = \langle 1,1;\ -1,1 \mid 0,\ 0\rangle = 1/\sqrt{3}$ and $\langle 1,1;\ 0,0 \mid 0,\ 0\rangle = -1/\sqrt{3}$.

Note that while the quintuplet states $|2,\ m\rangle$ (with $m = \pm 2, \pm 1, 0$) and the singlet state $|\ 0,\ 0\rangle$ are symmetric, the triplet states $|1,\ m\rangle$ (with $m = \pm 1, 0$) are antisymmetric under space inversion.

Problem 7.4

(a) Find the total spin of a system of three spin $\frac{1}{2}$ particles and derive the corresponding Clebsch–Gordan coefficients.

(b) Consider a system of three nonidentical spin $\frac{1}{2}$ particles whose Hamiltonian is given by $\hat{H} = -\epsilon_0(\hat{\vec{S}}_1 \cdot \hat{\vec{S}}_3 + \hat{\vec{S}}_2 \cdot \hat{\vec{S}}_3)/\hbar^2$. Find the system's energy levels and their degeneracies.

Solution

(a) To add $j_1 = \frac{1}{2}$, $j_2 = \frac{1}{2}$, and $j_3 = \frac{1}{2}$, we begin by coupling j_1 and j_2 to form $j_{12} = j_1 + j_2$, where $|j_1 - j_2| \le j_{12} \le j_1 + j_2$; hence $j_{12} = 0,\ 1$. Then we add j_{12} and j_3; this leads to $|j_{12} - j_3| \le j \le j_{12} + j_3$ or $j = \frac{1}{2},\ \frac{3}{2}$.

We are going to denote the joint eigenstates of $\hat{J}_1^2$, $\hat{J}_2^2$, $\hat{J}_3^2$, $\hat{J}_{12}^2$, $\hat{J}^2$, and $\hat{J}_z$ by $|j_{12}, j, m\rangle$ and the joint eigenstates of $\hat{J}_1^2$, $\hat{J}_2^2$, $\hat{J}_3^2$, $\hat{J}_{1_z}$, $\hat{J}_{2_z}$, and $\hat{J}_{3_z}$ by $|\ j_1, j_2, j_3;\ m_1, m_2, m_3\rangle$; since $j_1 = j_2 = j_3 = \frac{1}{2}$ and $m_1 = \pm\frac{1}{2}$, $m_2 = \pm\frac{1}{2}$, $m_3 = \pm\frac{1}{2}$, we will be using throughout this problem the lighter notation $|j_1, j_2, j_3; \pm, \pm, \pm\rangle$ to abbreviate $|\ \frac{1}{2}, \frac{1}{2}, \frac{1}{2}; \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}\rangle$.

In total there are eight states $|j_{12}, j, m\rangle$ since $(2j_1 + 1)(2j_2 + 1)(2j_3 + 1) = 8$. Four of these correspond to the subspace $j = \frac{3}{2}$: $|\ 1, \frac{3}{2}, \frac{3}{2}\rangle$, $|\ 1, \frac{3}{2}, \frac{1}{2}\rangle$, $|\ 1, \frac{3}{2}, -\frac{1}{2}\rangle$, and $|\ 1, \frac{3}{2}, -\frac{3}{2}\rangle$. The remaining four belong to the subspace $j = \frac{1}{2}$: $|\ 0, \frac{1}{2}, \frac{1}{2}\rangle$, $|\ 0, \frac{1}{2}, -\frac{1}{2}\rangle$, $|\ 1, \frac{1}{2}, \frac{1}{2}\rangle$, and $|\ 1, \frac{1}{2}, -\frac{1}{2}\rangle$. To construct the states $|j_{12}, j, m\rangle$ in terms of $|j_1, j_2, j_3; \pm, \pm, \pm\rangle$, we are going to consider the two subspaces $j = \frac{3}{2}$ and $j = \frac{1}{2}$ separately.

Subspace $j = \frac{3}{2}$

First, the states $|\,1,\frac{3}{2},\frac{3}{2}\rangle$ and $|\,1,\frac{3}{2},-\frac{3}{2}\rangle$ are clearly given by

$$\left|1,\frac{3}{2},\frac{3}{2}\right\rangle = |j_1,j_2,j_3;+,+,+\rangle, \qquad \left|1,\frac{3}{2},-\frac{3}{2}\right\rangle = |j_1,j_2,j_3;-,-,-\rangle. \tag{7.382}$$

To obtain $|\,1,\frac{3}{2},\frac{1}{2}\rangle$, we need to apply, on the one hand, $\hat{J}_-$ on $|\,1,\frac{3}{2},\frac{3}{2}\rangle$ (see (7.229)),

$$\hat{J}_-\left|1,\frac{3}{2},\frac{3}{2}\right\rangle = \hbar\sqrt{\frac{3}{2}\left(\frac{3}{2}+1\right)-\frac{3}{2}\left(\frac{3}{2}-1\right)}\left|1,\frac{3}{2},\frac{1}{2}\right\rangle = \hbar\sqrt{3}\left|1,\frac{3}{2},\frac{1}{2}\right\rangle, \tag{7.383}$$

and, on the other hand, apply $(\hat{J}_{1-}+\hat{J}_{2-}+\hat{J}_{3-})$ on $|j_1,j_2,j_3;+,+,+\rangle$ (see (7.230) to (7.232)). This yields

$$\begin{aligned}(\hat{J}_{1-}+\hat{J}_{2-}+\hat{J}_{3-})|j_1,j_2,j_3;+,+,+\rangle = \hbar\Big(&|j_1,j_2,j_3;-,+,+\rangle+|j_1,j_2,j_3;+,-,+\rangle\\ &+|j_1,j_2,j_3;+,+,-\rangle\Big),\end{aligned} \tag{7.384}$$

since $\sqrt{\frac{1}{2}(\frac{1}{2}+1)-\frac{1}{2}(\frac{1}{2}-1)}=1$. Equating (7.383) and (7.384) we infer

$$\left|1,\frac{3}{2},\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\Big(|j_1,j_2,j_3;-,+,+\rangle+|j_1,j_2,j_3;+,-,+\rangle+|j_1,j_2,j_3;+,+,-\rangle\Big). \tag{7.385}$$

Following the same method—applying $\hat{J}_-$ on $|\,1,\frac{3}{2},\frac{1}{2}\rangle$ and $(\hat{J}_{1-}+\hat{J}_{2-}+\hat{J}_{3-})$ on the right-hand side of (7.385) and then equating the two results—we find

$$\left|1,\frac{3}{2},-\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\Big(|j_1,j_2,j_3;+,-,-\rangle+|j_1,j_2,j_3;-,+,-\rangle+|j_1,j_2,j_3;-,-,+\rangle\Big). \tag{7.386}$$

Subspace $j=\frac{1}{2}$

We can write $|\,0,\frac{1}{2},\frac{1}{2}\rangle$ as a linear combination of $|j_1,j_2,j_3;+,-,+\rangle$ and $|j_1,j_2,j_3;-,+,+\rangle$:

$$\left|0,\frac{1}{2},\frac{1}{2}\right\rangle = \alpha|j_1,j_2,j_3;+,-,+\rangle+\beta|j_1,j_2,j_3;-,+,+\rangle. \tag{7.387}$$

Since $|\,0,\frac{1}{2},\frac{1}{2}\rangle$ is normalized, while $|j_1,j_2,j_3;+,-,+\rangle$ and $|j_1,j_2,j_3;-,+,+\rangle$ are orthonormal, and since the Clebsch–Gordan coefficients, such as α and β, are real numbers, equation (7.387) yields

$$\alpha^2+\beta^2=1. \tag{7.388}$$

On the other hand, since $\langle 1,\frac{3}{2},\frac{1}{2}\,|\,0,\frac{1}{2},\frac{1}{2}\rangle=0$, a combination of (7.385) and (7.387) leads to

$$\frac{1}{\sqrt{3}}(\alpha+\beta)=0 \quad\Longrightarrow\quad \alpha=-\beta. \tag{7.389}$$

A substitution of $\alpha=-\beta$ into (7.388) yields $\alpha=-\beta=1/\sqrt{2}$, and substituting this into (7.387) we obtain

$$\left|0,\frac{1}{2},\frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}}\Big(|j_1,j_2,j_3;+,-,+\rangle-|j_1,j_2,j_3;-,+,+\rangle\Big). \tag{7.390}$$

Following the same procedure that led to (7.385)—applying $\hat{J}_-$ on the left-hand side of (7.390) and $(\hat{J}_{1-} + \hat{J}_{2-} + \hat{J}_{3-})$ on the right-hand side and then equating the two results—we find

$$\left|0, \frac{1}{2}, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}}\left(|j_1, j_2, j_3; +, -, -\rangle - |j_1, j_2, j_3; -, +, -\rangle\right). \tag{7.391}$$

Now, to find $|\, 1, \frac{1}{2}, \frac{1}{2}\rangle$, we may write it as a linear combination of $|j_1, j_2, j_3; +, +, -\rangle$, $|j_1, j_2, j_3; +, -, +\rangle$, and $|j_1, j_2, j_3; -, +, +\rangle$:

$$\left|1, \frac{1}{2}, \frac{1}{2}\right\rangle = \alpha|j_1, j_2, j_3; +, +, -\rangle + \beta|j_1, j_2, j_3; +, -, +\rangle + \gamma|j_1, j_2, j_3; -, +, +\rangle. \tag{7.392}$$

This state is orthogonal to $|\, 0, \frac{1}{2}, \frac{1}{2}\rangle$, and hence $\beta = \gamma$; similarly, since this state is also orthogonal to $|\, 1, \frac{3}{2}, \frac{1}{2}\rangle$, we have $\alpha + \beta + \gamma = 0$, and hence $\alpha + 2\beta = 0$ or $\alpha = -2\beta = -2\gamma$. Now, since all the states of (7.392) are orthonormal, we have $\alpha^2 + \beta^2 + \gamma^2 = 1$, which when combined with $\alpha = -2\beta = -2\gamma$ leads to $\alpha = 2/\sqrt{6}$ and $\beta = \gamma = -1/\sqrt{6}$. We may thus write (7.392) as

$$\left|1, \frac{1}{2}, \frac{1}{2}\right\rangle = \frac{1}{\sqrt{6}}\left(2|j_1, j_2, j_3; +, +, -\rangle - |j_1, j_2, j_3; +, -, +\rangle - |j_1, j_2, j_3; -, +, +\rangle\right). \tag{7.393}$$

Finally, applying $\hat{J}_-$ on the left-hand side of (7.393) and $(\hat{J}_{1-} + \hat{J}_{2-} + \hat{J}_{3-})$ on the right-hand side and equating the two results, we find

$$\left|1, \frac{1}{2}, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{6}}\left(|j_1, j_2, j_3; -, +, -\rangle + |j_1, j_2, j_3; +, -, -\rangle - 2|j_1, j_2, j_3; -, -, +\rangle\right). \tag{7.394}$$

(b) Since we have three different (nonidentical) particles, their spin angular momenta mutually commute. We may thus write their Hamiltonian as $\hat{H} = -(\epsilon_0/\hbar^2)(\hat{\vec{S}}_1 + \hat{\vec{S}}_2) \cdot \hat{\vec{S}}_3$. Due to this suggestive form of $\hat{H}$, it is appropriate, as shown in (a), to start by coupling $\hat{\vec{S}}_1$ with $\hat{\vec{S}}_2$ to obtain $\hat{\vec{S}}_{12} = \hat{\vec{S}}_1 + \hat{\vec{S}}_2$, and then add $\hat{\vec{S}}_{12}$ to $\hat{\vec{S}}_3$ to generate the total spin: $\hat{\vec{S}} = \hat{\vec{S}}_{12} + \hat{\vec{S}}_3$. We may thus write $\hat{H}$ as

$$\hat{H} = -\frac{\epsilon_0}{\hbar^2}\left(\hat{\vec{S}}_1 + \hat{\vec{S}}_2\right) \cdot \hat{\vec{S}}_3 = -\frac{\epsilon_0}{\hbar^2}\hat{\vec{S}}_{12} \cdot \hat{\vec{S}}_3 = -\frac{\epsilon_0}{2\hbar^2}\left(\hat{S}^2 - \hat{S}_{12}^2 - \hat{S}_3^2\right), \tag{7.395}$$

since $\hat{\vec{S}}_{12} \cdot \hat{\vec{S}}_3 = \frac{1}{2}[(\hat{\vec{S}}_{12} + \hat{\vec{S}}_3)^2 - \hat{S}_{12}^2 - \hat{S}_3^2]$. Since the operators $\hat{H}$, $\hat{S}^2$, $\hat{S}_{12}^2$, and $\hat{S}_3^2$ mutually commute, we may select as their joint eigenstates the kets $|s_{12}, s, m\rangle$; we have seen in (a) how to construct these states. The eigenvalues of $\hat{H}$ are thus given by

$$\begin{aligned} \hat{H}|s_{12}, s, m\rangle &= -\frac{\epsilon_0}{2\hbar^2}\left(\hat{S}^2 - \hat{S}_{12}^2 - \hat{S}_3^2\right)|s_{12}, s, m\rangle \\ &= -\frac{\epsilon_0}{2}\left[s(s+1) - s_{12}(s_{12}+1) - \frac{3}{4}\right]|s_{12}, s, m\rangle, \end{aligned} \tag{7.396}$$

since $s_3 = \frac{1}{2}$ and $\hat{S}_3^2|s_{12}, s, m\rangle = \hbar^2 s_3(s_3+1)|s_{12}, s, m\rangle = (3\hbar^2/4)|s_{12}, s, m\rangle$.

As shown in (7.396), the energy levels of this system are degenerate with respect to m, since they depend on the quantum numbers s and s_{12} but not on m:

$$E_{s_{12},s} = -\frac{\epsilon_0}{2}\left[s(s+1) - s_{12}(s_{12}+1) - \frac{3}{4}\right]. \tag{7.397}$$

For instance, the energy $E_{s_{12},s} = E_{1,3/2} = -\epsilon_0/2$ is fourfold degenerate, since it corresponds to four different states: $|s_{12},\ s,\ m\rangle = |\ 1, \frac{3}{2}, \pm\frac{3}{2}\rangle$ and $|\ 1, \frac{3}{2}, \pm\frac{1}{2}\rangle$. Similarly, the energy $E_{0,1/2} = 0$ is twofold degenerate; the corresponding states are $|\ 0, \frac{1}{2}, \pm\frac{1}{2}\rangle$. Finally, the energy $E_{1,1/2} = \epsilon_0$ is also twofold degenerate since it corresponds to $|\ 1, \frac{1}{2}, \pm\frac{1}{2}\rangle$.

Problem 7.5
Consider a system of four nonidentical spin $\frac{1}{2}$ particles. Find the possible values of the total spin S of this system and specify the number of angular momentum eigenstates, corresponding to each value of S.

Solution
First, we need to couple two spins at a time: $\hat{\vec{S}}_{12} = \hat{\vec{S}}_1 + \hat{\vec{S}}_2$ and $\hat{\vec{S}}_{34} = \hat{\vec{S}}_3 + \hat{\vec{S}}_4$. Then we couple $\hat{\vec{S}}_{12}$ and $\hat{\vec{S}}_{34}$: $\hat{\vec{S}} = \hat{\vec{S}}_{12} + \hat{\vec{S}}_{34}$. From Problem 7.4, page 484, we have $s_{12} = 0, 1$ and $s_{34} = 0, 1$. In total there are 16 states $|sm\rangle$ since $(2s_1 + 1)(2s_2 + 1)(2s_3 + 1)(2s_4 + 1) = 2^4 = 16$.

Since $s_{12} = 0, 1$ and $s_{34} = 0, 1$, the coupling of $\hat{\vec{S}}_{12}$ and $\hat{\vec{S}}_{34}$ yields the following values for the total spin s:

- When $s_{12} = 0$ and $s_{34} = 0$ we have only one possible value, $s = 0$, and hence only one eigenstate, $|sm\rangle = |\ 0,\ 0\rangle$.
- When $s_{12} = 1$ and $s_{34} = 0$, we have $s = 1$; there are three eigenstates: $|s\ m\rangle = |\ 1,\ \pm 1\rangle$, and $|\ 1,\ 0\rangle$.
- When $s_{12} = 0$ and $s_{34} = 1$, we have $s = 1$; there are three eigenstates: $|sm\rangle = |\ 1,\ \pm 1\rangle$, and $|\ 1,\ 0\rangle$.
- When $s_{12} = 1$ and $s_{34} = 1$ we have $s = 0, 1, 2$; we have here nine eigenstates (see Problem 7.3, page 481): $|\ 0,\ 0\rangle$, $|\ 1,\ \pm 1\rangle$, $|\ 1,\ 0\rangle$, $|\ 2,\ \pm 2\rangle$, $|\ 2,\ \pm 1\rangle$, and $|\ 2,\ 0\rangle$.

In conclusion, the possible values of the total spin when coupling four $\frac{1}{2}$ spins are $s = 0,\ 1,\ 2$; the value $s = 0$ occurs twice, $s = 1$ three times, and $s = 2$ only once.

Problem 7.6
Work out the coupling of the isospins of a pion–nucleon system and infer the various states of this system.

Solution
Since the isospin of a pion meson is 1 and that of a nucleon is $\frac{1}{2}$, the total isospin of a pion–nucleon system can be obtained by coupling the isospins $t_1 = 1$ and $t_2 = \frac{1}{2}$. The various values of the total isospin lie in the range $|t_1 - t_2| < T < t_1 + t_2$; hence they are given by $T = \frac{3}{2}, \frac{1}{2}$.

The coupling of the isospins $t_1 = 1$ and $t_2 = \frac{1}{2}$ is analogous to the addition of an orbital angular momentum $l = 1$ and a spin $\frac{1}{2}$; the expressions pertaining to this coupling are listed in (7.215) to (7.220). Note that there are three different π-mesons:

$$|\ 1,\ 1\rangle = |\ \pi^+\rangle, \qquad |\ 1,\ 0\rangle = |\ \pi^0\rangle, \qquad |\ 1,\ -1\rangle = |\ \pi^-\rangle, \tag{7.398}$$

and two nucleons, a proton and a neutron:

$$\left|\frac{1}{2},\ \frac{1}{2}\right\rangle = |\ p\rangle, \qquad \left|\frac{1}{2},\ -\frac{1}{2}\right\rangle = |\ n\rangle. \tag{7.399}$$

By analogy with (7.215) to (7.220) we can write the states corresponding to $T = \frac{3}{2}$ as

$$\left|\frac{3}{2}, \frac{3}{2}\right\rangle = |1, 1\rangle\left|\frac{1}{2}, \frac{1}{2}\right\rangle = |\pi^+\rangle\,|p\rangle, \tag{7.400}$$

$$\left|\frac{3}{2}, \frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\,|1, 0\rangle\left|\frac{1}{2}, \frac{1}{2}\right\rangle + \frac{1}{\sqrt{3}}\,|1, 1\rangle\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\,|\pi^0\rangle\,|p\rangle + \frac{1}{\sqrt{3}}\,|\pi^+\rangle\,|n\rangle, \tag{7.401}$$

$$\left|\frac{3}{2}, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\,|1,-1\rangle\left|\frac{1}{2}, \frac{1}{2}\right\rangle + \sqrt{\frac{2}{3}}\,|1,0\rangle\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\,|\pi^-\rangle\,|p\rangle + \sqrt{\frac{2}{3}}\,|\pi^0\rangle\,|n\rangle, \tag{7.402}$$

$$\left|\frac{3}{2}, -\frac{3}{2}\right\rangle = |1, -1\rangle\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = |\pi^-\rangle\,|n\rangle, \tag{7.403}$$

and those corresponding to $T = \frac{1}{2}$ as

$$\left|\frac{1}{2}, \frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\,|1, 1\rangle\left|\frac{1}{2}, -\frac{1}{2}\right\rangle - \frac{1}{\sqrt{3}}\,|1, 0\rangle\left|\frac{1}{2}, \frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\,|\pi^+\rangle\,|n\rangle - \frac{1}{\sqrt{3}}\,|\pi^0\rangle\,|p\rangle, \tag{7.404}$$

$$\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\,|1,0\rangle\left|\frac{1}{2}, -\frac{1}{2}\right\rangle - \sqrt{\frac{2}{3}}\,|1, -1\rangle\left|\frac{1}{2}, \frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\,|\pi^0\rangle\,|n\rangle - \sqrt{\frac{2}{3}}\,|\pi^-\rangle\,|p\rangle. \tag{7.405}$$

Problem 7.7

(a) Calculate the expression of $\langle 2, 0|Y_{10}\,|\,1, 0\rangle$.

(b) Use the result of (a) along with the Wigner–Eckart theorem to calculate the reduced matrix element $\langle 2 \,\|\, Y_1 \,\|\, 1\rangle$.

Solution

(a) Since

$$\langle 2, 0\,|\,Y_{10}\,|\,1, 0\rangle = \int_0^{\pi} \sin\theta\, d\theta \int_0^{2\pi} Y_{20}^*(\theta,\varphi) Y_{10}(\theta,\varphi) Y_{10}(\theta,\varphi)\, d\varphi, \tag{7.406}$$

and using the relations $Y_{20}(\theta,\varphi) = \sqrt{5/(16\pi)}(3\cos^2\theta - 1)$ and $Y_{10}(\theta,\varphi) = \sqrt{3/(4\pi)}\cos\theta$, we have

$$\begin{aligned}\langle 2, 0\,|\,Y_{10}\,|\,1, 0\rangle &= \frac{3}{4\pi}\sqrt{\frac{5}{16\pi}} \int_0^{\pi} \cos^2\theta(3\cos^2\theta - 1)\sin\theta\, d\theta \int_0^{2\pi} d\varphi \\ &= \frac{3}{2}\sqrt{\frac{5}{16\pi}} \int_0^{\pi} \cos^2\theta(3\cos^2\theta - 1)\sin\theta\, d\theta. \end{aligned} \tag{7.407}$$

The change of variables $x = \cos\theta$ leads to

$$\langle 2,\ 0 \mid Y_{10} \mid 1,\ 0\rangle = \frac{3}{2}\sqrt{\frac{5}{16\pi}}\int_0^{\pi} \cos^2\theta(3\cos^3\theta - 1)\sin\theta\, d\theta$$
$$= \frac{3}{2}\sqrt{\frac{5}{16\pi}}\int_{-1}^{1} x^2(3x^2 - 1)\, dx = \frac{1}{\sqrt{5\pi}}. \tag{7.408}$$

(b) Applying the Wigner–Eckart theorem to $\langle 2,\ 0 \mid Y_{10} \mid 1,\ 0\rangle$ and using the Clebsch–Gordan coefficient $\langle 1,1;\ 0,0 \mid 2,\ 0\rangle = 2/\sqrt{6}$, we have

$$\langle 2,\ 0 \mid Y_{10} \mid 1,\ 0\rangle = \langle 1,1;\ 0,0 \mid 2,\ 0\rangle\langle 2 \parallel Y_1 \parallel 1\rangle = \frac{2}{\sqrt{6}}\langle 2 \parallel Y_1 \parallel 1\rangle. \tag{7.409}$$

Finally, we may obtain $\langle 2 \parallel Y_1 \parallel 1\rangle$ from (7.408) and (7.409):

$$\langle 2 \parallel Y_1 \parallel 1\rangle = \sqrt{\frac{3}{10\pi}}. \tag{7.410}$$

Problem 7.8

(a) Find the reduced matrix elements associated with the spherical harmonic $Y_{kq}(\theta,\varphi)$.
(b) Calculate the dipole transitions $\langle n'l'm' \mid \vec{r} \mid nlm\rangle$.

Solution

On the one hand, an application of the Wigner–Eckart theorem to Y_{kq} yields

$$\langle l',\ m' \mid Y_{kq} \mid l,\ m\rangle = \langle l,k;\ m,q \mid l',\ m'\rangle\langle l' \parallel Y^{(k)} \parallel l\rangle \tag{7.411}$$

and, on the other hand, a straightforward evaluation of

$$\langle l',\ m' \mid Y_{kq} \mid l,\ m\rangle = \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta\, d\theta\langle l',\ m \mid \theta\varphi\rangle Y_{kq}(\theta,\varphi)\langle\theta\varphi \mid l,\ m\rangle$$
$$= \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta Y^*_{l'm'}(\theta,\varphi)Y_{kq}(\theta,\varphi)Y_{lm}(\theta,\varphi) \tag{7.412}$$

can be inferred from the triple integral relation (7.253):

$$\boxed{\langle l',\ m' \mid Y_{kq} \mid l,\ m\rangle = \sqrt{\frac{(2l+1)(2k+1)}{4\pi(2l'+1)}}\langle l,k;\ 0,0|l',\ 0\rangle\langle l,k;\ m,q \mid l',\ m'\rangle.} \tag{7.413}$$

We can then combine (7.411) and (7.413) to obtain the reduced matrix element

$$\boxed{\langle l' \parallel Y^{(k)} \parallel l\rangle = \sqrt{\frac{(2l+1)(2k+1)}{4\pi(2l'+1)}}\langle l,k;\ 0,0|l',\ 0\rangle.} \tag{7.414}$$

(b) To calculate $\langle n'l'm' \mid \vec{r} \mid nlm\rangle$ it is more convenient to express the vector $\vec{r}$ in terms of the spherical components $\vec{r} = (r_{-1}, r_0, r_1)$, which are given in terms of the Cartesian coordinates x, y, z as follows:

$$r_1 = -\frac{x+iy}{\sqrt{2}} = \frac{r}{\sqrt{2}}e^{i\varphi}\sin\theta, \qquad r_0 = z = r\cos\theta, \qquad r_{-1} = \frac{x-iy}{\sqrt{2}} = \frac{r}{\sqrt{2}}e^{-i\varphi}\sin\theta, \tag{7.415}$$

which in turn may be condensed into a single relation

$$r_q = \sqrt{\frac{4\pi}{3}}rY_{1q}(\theta,\varphi), \qquad q = 1, 0, -1. \tag{7.416}$$

Next we may write $\langle n'l'm' \mid r_q \mid nlm\rangle$ in terms of a radial part and an angular part:

$$\langle n'l'm' \mid r_q \mid nlm\rangle = \sqrt{\frac{4\pi}{3}}\langle n'l' \mid r_q|nl\rangle\langle l',\ m'|Y_{1q}(\theta,\varphi) \mid l,\ m\rangle. \tag{7.417}$$

The calculation of the radial part, $\langle n'l' \mid r_q \mid nl\rangle = \int_0^\infty r^3 R^*_{n'l'}(r)R_{nl}(r)\,dr$, is straightforward and is of no concern to us here; see Chapter 6 for its calculation. As for the angular part $\langle l',\ m' \mid Y_{1q}(\theta,\varphi) \mid l,\ m\rangle$, we can infer its expression from (7.413)

$$\langle l',\ m' \mid Y_{1q} \mid l,\ m\rangle = \sqrt{\frac{3(2l+1)}{4\pi(2l'+1)}}\langle l, 1;\ 0, 0|l',\ 0\rangle\langle l, 1;\ m, q \mid l',\ m'\rangle. \tag{7.418}$$

The Clebsch–Gordan coefficients $\langle l, 1;\ m, q \mid l',\ m'\rangle$ vanish unless $m' = m+q$ and $l-1 \le l' \le l+1$ or $\Delta m = m' - m = q = 1, 0, -1$ and $\Delta l = l' - l = 1, 0, -1$. Notice that the case $\Delta l = 0$ is ruled out from the parity selection rule; so, the only permissible values of l' and l are those for which $\Delta l = l' - l = \pm 1$. Obtaining the various relevant Clebsch–Gordan coefficients from standard tables, we can ascertain that the only terms of (7.418) that survive are

$$\langle l+1,\ m+1|Y_{11} \mid l,\ m\rangle = \sqrt{\frac{3(l+m+1)(l+m+2)}{8\pi(2l+1)(2l+3)}}, \tag{7.419}$$

$$\langle l-1,\ m+1|Y_{11} \mid l,\ m\rangle = \sqrt{\frac{3(l-m-1)(l-m)}{8\pi(2l+1)(2l+3)}}, \tag{7.420}$$

$$\langle l+1,\ m|Y_{10} \mid l,\ m\rangle = \sqrt{\frac{3[(l+1)^2-m^2]}{4\pi(2l+1)(2l+3)}}, \tag{7.421}$$

$$\langle l-1,\ m|Y_{10} \mid l,\ m\rangle = \sqrt{\frac{3(l^2-m^2)}{4\pi(2l+1)(2l-1)}}, \tag{7.422}$$

$$\langle l+1,\ m-1|Y_{1-1} \mid l,\ m\rangle = \sqrt{\frac{3(l-m+1)(l-m+2)}{8\pi(2l+1)(2l+3)}}, \tag{7.423}$$

$$\langle l-1,\ m-1|Y_{1-1} \mid l,\ m\rangle = \sqrt{\frac{3(l+m)(l+m-1)}{8\pi(2l+1)(2l-1)}}. \tag{7.424}$$

Problem 7.9
Find the rotation matrix $d^{(1)}$ corresponding to $j = 1$.

Solution
To find the matrix of $d^{(1)}(\beta) = e^{-i\beta \hat{J}_y/\hbar}$ for $j = 1$, we need first to find the matrix representation of $\hat{J}_y$ within the joint eigenstates $\{|\, j,\, m\rangle\}$ of $\hat{J}^2$ and $\hat{J}_z$. Since the basis of $j = 1$ consists of three states $|\, 1,\, -1\rangle, |\, 1,\, 0\rangle, |\, 1,\, 1\rangle$, the matrix representing $\hat{J}_y$ within this basis is given by

$$J_y = \frac{\hbar}{2}\begin{pmatrix} \langle 1,\,1\,|\,\hat{J}_y\,|\,1,\,1\rangle & \langle 1,\,1\,|\,\hat{J}_y\,|\,1,\,0\rangle & \langle 1,\,1\,|\,\hat{J}_y\,|\,1,\,-1\rangle \\ \langle 1,\,0\,|\,\hat{J}_y\,|\,1,\,1\rangle & \langle 1,\,0\,|\,\hat{J}_y\,|\,1,\,0\rangle & \langle 1,\,0\,|\,\hat{J}_y\,|\,1,\,-1\rangle \\ \langle 1,\,-1\,|\,\hat{J}_y\,|\,1,\,1\rangle & \langle 1,\,-1\,|\,\hat{J}_y\,|\,1,\,0\rangle & \langle 1,\,-1\,|\,\hat{J}_y\,|\,1,\,-1\rangle \end{pmatrix}$$
$$= \frac{i\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{7.425}$$

We can easily verify that $J_y^3 = J_y$:

$$J_y^2 = \frac{\hbar^2}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \qquad J_y^3 = \frac{i\hbar^3}{\sqrt{2}}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \hbar^2 J_y. \tag{7.426}$$

We can thus infer

$$J_y^{2n} = \hbar^{2n-2} J_y^2 \qquad (n > 0), \qquad\qquad J_y^{2n+1} = \hbar^{2n} J_y. \tag{7.427}$$

Combining these two relations with

$$e^{-i\beta \hat{J}_y/\hbar} = \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{i\beta}{\hbar}\right)^n J_y^n$$
$$= \sum_{n=0}^{\infty} \frac{1}{(2n)!}\left(-\frac{i\beta}{\hbar}\right)^{2n} J_y^{2n} + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!}\left(-\frac{i\beta}{\hbar}\right)^{2n+1} J_y^{2n+1}, \tag{7.428}$$

we obtain

$$e^{-i\beta \hat{J}_y/\hbar} = \hat{I} + \left(\frac{\hat{J}_y}{\hbar}\right)^2 \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!}(\beta)^{2n} - i\frac{\hat{J}_y}{\hbar}\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\beta^{2n+1}$$
$$= \hat{I} + \left(\frac{\hat{J}_y}{\hbar}\right)^2 \left[\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}(\beta)^{2n} - 1\right] - i\frac{\hat{J}_y}{\hbar}\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\beta^{2n+1}, \tag{7.429}$$

where $\hat{I}$ is the 3×3 unit matrix. Using the relations $\sum_{n=0}^{\infty}[(-1)^n/(2n)!](\beta)^{2n} = \cos\beta$ and $\sum_{n=0}^{\infty}[(-1)^n/(2n+1)!]\beta^{2n+1} = \sin\beta$, we may write

$$e^{-i\beta \hat{J}_y/\hbar} = \hat{I} + \left(\frac{\hat{J}_y}{\hbar}\right)^2 [\cos\beta - 1] - i\frac{\hat{J}_y}{\hbar}\sin\beta. \tag{7.430}$$

Inserting now the matrix expressions for J_y and J_y^2 as listed in (7.425) and (7.426), we obtain

$$e^{-i\beta \hat{J}_y/\hbar} = \hat{I} + \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix}(\cos\beta - 1) - i\frac{i}{\sqrt{2}}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}\sin\beta \tag{7.431}$$

or

$$d^{(1)}(\beta) = \begin{pmatrix} d^{(1)}_{11} & d^{(1)}_{1,0} & d^{(1)}_{1-1} \\ d^{(1)}_{01} & d^{(1)}_{00} & d^{(1)}_{0-1} \\ d^{(1)}_{-11} & d^{(1)}_{-10} & d^{(1)}_{-1-1} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(1+\cos\beta) & -\frac{1}{\sqrt{2}}\sin\beta & \frac{1}{2}(1-\cos\beta) \\ \frac{1}{\sqrt{2}}\sin\beta & \cos\beta & -\frac{1}{\sqrt{2}}\sin\beta \\ \frac{1}{2}(1-\cos\beta) & \frac{1}{\sqrt{2}}\sin\beta & \frac{1}{2}(1+\cos\beta) \end{pmatrix}. \tag{7.432}$$

Since $\frac{1}{2}(1+\cos\beta) = \cos^2(\beta/2)$ and $\frac{1}{2}(1-\cos\beta) = \sin^2(\beta/2)$, we have

$$d^{(1)}(\beta) = e^{-i\beta J_y/\hbar} = \begin{pmatrix} \cos^2(\beta/2) & -\frac{1}{\sqrt{2}}\sin(\beta) & \sin^2(\beta/2) \\ \frac{1}{\sqrt{2}}\sin(\beta) & \cos(\beta) & -\frac{1}{\sqrt{2}}\sin(\beta) \\ \sin^2(\beta/2) & \frac{1}{\sqrt{2}}\sin(\beta) & \cos^2(\beta/2) \end{pmatrix}. \tag{7.433}$$

This method becomes quite intractable when attempting to derive the matrix of $d^{(j)}(\beta)$ for large values of j. In Problem 7.10 we are going to present a simpler method for deriving $d^{(j)}(\beta)$ for larger values of j; this method is based on the addition of angular momenta.

Problem 7.10

(a) Use the relation

$$d^{(j)}_{mm'}(\beta) = \sum_{m_1 m_2}\sum_{m_1' m_2'} \langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle\langle j_1, j_2;\ m_1', m_2' \mid j,\ m'\rangle\, d^{(j_1)}_{m_1 m_1'}(\beta) d^{(j_2)}_{m_2 m_2'}(\beta),$$

for the case where $j_1 = 1$ and $j_2 = \frac{1}{2}$ along with the Clebsch–Gordan coefficients derived in (7.215) to (7.218), and the matrix elements of $d^{(1/2)}(\beta)$ and $d^{(1)}(\beta)$, which are given by (7.97) and (7.433), respectively, to find the expressions of the matrix elements of $d^{(3/2)}_{\frac{3}{2}\frac{3}{2}}(\beta)$, $d^{(3/2)}_{\frac{3}{2}\frac{1}{2}}(\beta)$, $d^{(3/2)}_{\frac{3}{2}-\frac{1}{2}}(\beta)$, $d^{(3/2)}_{\frac{3}{2}-\frac{3}{2}}(\beta)$, $d^{(3/2)}_{\frac{1}{2}\frac{1}{2}}(\beta)$, and $d^{(3/2)}_{\frac{1}{2}-\frac{1}{2}}(\beta)$.

(b) Use the six expressions derived in (a) to infer the matrix of $d^{(3/2)}(\beta)$.

Solution

(a) Using $\langle 1, \frac{1}{2};\ 1, \frac{1}{2} \mid \frac{3}{2},\ \frac{3}{2}\rangle = 1$, $d^{(1)}_{11}(\beta) = \cos^2(\beta/2)$ and $d^{(1/2)}_{\frac{1}{2}\frac{1}{2}}(\beta) = \cos(\beta/2)$, we have

$$d^{(3/2)}_{\frac{3}{2}\frac{3}{2}}(\beta) = \left\langle 1, \frac{1}{2};\ 1, \frac{1}{2}\middle|\frac{3}{2},\ \frac{3}{2}\right\rangle\left\langle 1, \frac{1}{2};\ 1, \frac{1}{2}\middle|\frac{3}{2},\ \frac{3}{2}\right\rangle d^{(1)}_{11}(\beta) d^{(1/2)}_{\frac{1}{2}\frac{1}{2}}(\beta) = \cos^3\left(\frac{\beta}{2}\right). \tag{7.434}$$

Similarly, since $\langle 1, \frac{1}{2};\ 0, \frac{1}{2} \mid \frac{3}{2},\ \frac{1}{2}\rangle = \sqrt{2/3}$, $\langle 1, \frac{1}{2};\ 1, -\frac{1}{2} \mid \frac{3}{2},\ \frac{1}{2}\rangle = 1/\sqrt{3}$, and since $d^{(1)}_{10}(\beta) =$

$-(1/\sqrt{2})\sin(\beta)$ and $d^{(1/2)}_{\frac{1}{2}-\frac{1}{2}}(\beta) = -\sin(\beta/2)$, we have

$$\begin{aligned} d^{(3/2)}_{\frac{3}{2}\frac{1}{2}}(\beta) &= \left\langle 1,\frac{1}{2};\,1,\frac{1}{2}\,\middle|\,\frac{3}{2},\,\frac{3}{2}\right\rangle\left\langle 1,\frac{1}{2};\,0,\frac{1}{2}\,\middle|\,\frac{3}{2},\,\frac{1}{2}\right\rangle d^{(1)}_{10}(\beta)d^{(1/2)}_{\frac{1}{2}\frac{1}{2}}(\beta) \\ &\quad + \left\langle 1,\frac{1}{2};\,1,\frac{1}{2}\,\middle|\,\frac{3}{2},\,\frac{3}{2}\right\rangle\left\langle 1,\frac{1}{2};\,1,-\frac{1}{2}\,\middle|\,\frac{3}{2},\,\frac{1}{2}\right\rangle d^{(1)}_{11}(\beta)d^{(1/2)}_{\frac{1}{2}-\frac{1}{2}}(\beta) \\ &= -\frac{1}{\sqrt{3}}\sin\beta\cos\left(\frac{\beta}{2}\right) - \frac{1}{\sqrt{3}}\cos^2\left(\frac{\beta}{2}\right)\sin\left(\frac{\beta}{2}\right) \\ &= -\sqrt{3}\sin\left(\frac{\beta}{2}\right)\cos^2\left(\frac{\beta}{2}\right). \end{aligned} \tag{7.435}$$

To calculate $d^{(3/2)}_{\frac{3}{2}-\frac{1}{2}}(\beta)$, we need to use the coefficients $\langle 1,\frac{1}{2};\,0,-\frac{1}{2}\mid\frac{3}{2},\,-\frac{1}{2}\rangle = \sqrt{2/3}$ and $\langle 1,\frac{1}{2};\,-1,\frac{1}{2}\mid\frac{3}{2},\,-\frac{1}{2}\rangle = 1/\sqrt{3}$ along with $d^{(1)}_{1,-1}(\beta) = \sin^2(\beta/2)$:

$$\begin{aligned} d^{(3/2)}_{\frac{3}{2}-\frac{1}{2}}(\beta) &= \left\langle 1,\frac{1}{2};\,1,\frac{1}{2}\,\middle|\,\frac{3}{2},\,\frac{3}{2}\right\rangle\left\langle 1,\frac{1}{2};\,-1,\frac{1}{2}\,\middle|\,\frac{3}{2},\,-\frac{1}{2}\right\rangle d^{(1)}_{1-1}(\beta)d^{(1/2)}_{\frac{1}{2}\frac{1}{2}}(\beta) \\ &\quad + \left\langle 1,\frac{1}{2};\,1,\frac{1}{2}\,\middle|\,\frac{3}{2},\,\frac{3}{2}\right\rangle\left\langle 1,\frac{1}{2};\,0,-\frac{1}{2}\,\middle|\,\frac{3}{2},\,-\frac{1}{2}\right\rangle d^{(1)}_{10}(\beta)d^{(1/2)}_{\frac{1}{2}-\frac{1}{2}}(\beta) \\ &= \frac{1}{\sqrt{3}}\sin^2\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right) + \frac{1}{\sqrt{3}}\sin\beta\sin\left(\frac{\beta}{2}\right) \\ &= \sqrt{3}\sin^2\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right). \end{aligned} \tag{7.436}$$

For $d^{(3/2)}_{\frac{3}{2}-\frac{3}{2}}(\beta)$ we have

$$d^{(3/2)}_{\frac{3}{2}-\frac{3}{2}}(\beta) = \left\langle 1,\frac{1}{2};1,\frac{1}{2}\,\middle|\,\frac{3}{2},\frac{3}{2}\right\rangle\left\langle 1,\frac{1}{2};\,-1,-\frac{1}{2}\,\middle|\,\frac{3}{2},\,-\frac{3}{2}\right\rangle d^{(1)}_{1-1}(\beta)d^{(1/2)}_{\frac{1}{2},-\frac{1}{2}}(\beta) = -\sin^3\left(\frac{\beta}{2}\right), \tag{7.437}$$

because $\langle 1,\frac{1}{2};\,-1,-\frac{1}{2}\mid\frac{3}{2},\,-\frac{3}{2}\rangle = 1$, $d^{(1)}_{1-1}(\beta) = \sin^2(\beta/2)$, and $d^{(1/2)}_{\frac{1}{2}-\frac{1}{2}}(\beta) = -\sin(\beta/2)$.

To calculate $d^{(3/2)}_{\frac{1}{2}\frac{1}{2}}(\beta)$, we need to use the coefficients $\langle 1,\frac{1}{2};\,0,\frac{1}{2}\mid\frac{3}{2},\,\frac{1}{2}\rangle = \sqrt{2/3}$ and

$\langle 1, \frac{1}{2};\ 1, -\frac{1}{2} \mid \frac{3}{2},\ \frac{1}{2}\rangle = 1/\sqrt{3}$ along with $d^{(1)}_{1-1}(\beta) = \sin^2(\beta/2)$:

$$\begin{aligned}
d^{(3/2)}_{\frac{1}{2}\frac{1}{2}}(\beta) &= \left\langle 1, \frac{1}{2};\ 1, -\frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle \left\langle 1, \frac{1}{2};\ 1, -\frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle d^{(1)}_{11}(\beta) d^{(1/2)}_{-\frac{1}{2}-\frac{1}{2}}(\beta) \\
&\quad + \left\langle 1, \frac{1}{2};\ 0, \frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle \left\langle 1, \frac{1}{2};\ 1, -\frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle d^{(1)}_{01}(\beta) d^{(1/2)}_{\frac{1}{2}-\frac{1}{2}}(\beta) \\
&\quad + \left\langle 1, \frac{1}{2};\ 0, \frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle \left\langle 1, \frac{1}{2};\ 0, \frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle d^{(1)}_{00}(\beta) d^{(1/2)}_{\frac{1}{2}\frac{1}{2}}(\beta) \\
&\quad + \left\langle 1, \frac{1}{2};\ 1, -\frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle \left\langle 1, \frac{1}{2};\ 0, \frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle d^{(1)}_{10}(\beta) d^{(1/2)}_{-\frac{1}{2}\frac{1}{2}}(\beta) \\
&= \frac{1}{3}\cos^3\left(\frac{\beta}{2}\right) - \frac{1}{3}\sin(\beta)\sin\left(\frac{\beta}{2}\right) + \frac{2}{3}\cos(\beta)\cos\left(\frac{\beta}{2}\right) - \frac{1}{3}\sin(\beta)\sin\left(\frac{\beta}{2}\right) \\
&= \left[3\cos^2\left(\frac{\beta}{2}\right) - 2\right]\cos\left(\frac{\beta}{2}\right) \\
&= \frac{1}{2}(3\cos\beta - 1)\cos\left(\frac{\beta}{2}\right). \qquad (7.438)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
d^{(3/2)}_{\frac{1}{2}-\frac{1}{2}}(\beta) &= \left\langle 1, \frac{1}{2};\ 1, -\frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle \left\langle 1, \frac{1}{2};\ -1, \frac{1}{2} \middle| \frac{3}{2},\ -\frac{1}{2}\right\rangle d^{(1)}_{1-1}(\beta) d^{(1/2)}_{-\frac{1}{2}\frac{1}{2}}(\beta) \\
&\quad + \left\langle 1, \frac{1}{2};\ 1, -\frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle \left\langle 1, \frac{1}{2};\ 0, -\frac{1}{2} \middle| \frac{3}{2},\ -\frac{1}{2}\right\rangle d^{(1)}_{10}(\beta) d^{(1/2)}_{-\frac{1}{2}-\frac{1}{2}}(\beta) \\
&\quad + \left\langle 1, \frac{1}{2};\ 0, \frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle \left\langle 1, \frac{1}{2};\ -1, \frac{1}{2} \middle| \frac{3}{2},\ -\frac{1}{2}\right\rangle d^{(1)}_{0-1}(\beta) d^{(1/2)}_{\frac{1}{2}\frac{1}{2}}(\beta) \\
&\quad + \left\langle 1, \frac{1}{2};\ 0, \frac{1}{2} \middle| \frac{3}{2},\ \frac{1}{2}\right\rangle \left\langle 1, \frac{1}{2};\ 0, -\frac{1}{2} \middle| \frac{3}{2},\ -\frac{1}{2}\right\rangle d^{(1)}_{00}(\beta) d^{(1/2)}_{\frac{1}{2}-\frac{1}{2}}(\beta) \\
&= \frac{1}{3}\sin^3\left(\frac{\beta}{2}\right) - \frac{1}{3}\sin(\beta)\cos\left(\frac{\beta}{2}\right) - \frac{1}{3}\sin(\beta)\cos\left(\frac{\beta}{2}\right) - \frac{2}{3}\cos(\beta)\sin\left(\frac{\beta}{2}\right) \\
&= -\left[3\cos^2\left(\frac{\beta}{2}\right) - 1\right]\sin\left(\frac{\beta}{2}\right) \\
&= -\frac{1}{2}(3\cos\beta + 1)\sin\left(\frac{\beta}{2}\right). \qquad (7.439)
\end{aligned}$$

(b) The remaining ten matrix elements of $d^{(3/2)}(\beta)$ can be inferred from the six elements derived above by making use of the properties of the d-function listed in (7.75). For instance, using $d^{(j)}_{m'm}(\beta) = (-1)^{m'-m} d^{(j)}_{-m'-m}(\beta)$, we can verify that

$$d^{(3/2)}_{-\frac{3}{2}-\frac{3}{2}}(\beta) = d^{(3/2)}_{\frac{3}{2}\frac{3}{2}}(\beta), \qquad d^{(3/2)}_{-\frac{1}{2}-\frac{1}{2}}(\beta) = d^{(3/2)}_{\frac{1}{2}\frac{1}{2}}(\beta), \qquad d^{(3/2)}_{-\frac{3}{2}-\frac{1}{2}}(\beta) = -d^{(3/2)}_{\frac{3}{2}\frac{1}{2}}(\beta), \qquad (7.440)$$

$$d^{(3/2)}_{-\frac{3}{2}\frac{1}{2}}(\beta) = d^{(3/2)}_{\frac{3}{2}-\frac{1}{2}}(\beta), \qquad d^{(3/2)}_{-\frac{3}{2}\frac{3}{2}}(\beta) = -d^{(3/2)}_{\frac{3}{2}-\frac{3}{2}}(\beta), \qquad d^{(3/2)}_{-\frac{1}{2}\frac{1}{2}}(\beta) = -d^{(3/2)}_{\frac{3}{2}-\frac{1}{2}}(\beta). \qquad (7.441)$$

Similarly, using $d^{(j)}_{m'm}(\beta) = (-1)^{m'-m} d^{(j)}_{mm'}(\beta)$ we can obtain the remaining four elements:

$$d^{(3/2)}_{\frac{1}{2}\frac{3}{2}}(\beta) = -d^{(3/2)}_{\frac{3}{2}\frac{1}{2}}(\beta), \qquad d^{(3/2)}_{-\frac{1}{2}\frac{3}{2}}(\beta) = d^{(3/2)}_{\frac{3}{2}-\frac{1}{2}}(\beta), \tag{7.442}$$

$$d^{(3/2)}_{\frac{1}{2}-\frac{3}{2}}(\beta) = d^{(3/2)}_{-\frac{3}{2}\frac{1}{2}}(\beta), \qquad d^{(3/2)}_{-\frac{1}{2}-\frac{3}{2}}(\beta) = -d^{(3/2)}_{-\frac{3}{2}-\frac{1}{2}}(\beta). \tag{7.443}$$

Collecting the six matrix elements calculated in (a) along with the ten elements inferred above, we obtain the matrix of $d^{(3/2)}(\beta)$:

$$\begin{pmatrix} \cos^3\left(\frac{\beta}{2}\right) & -\sqrt{3}\sin\left(\frac{\beta}{2}\right)\cos^2\left(\frac{\beta}{2}\right) & \sqrt{3}\sin^2\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right) & -\sin^3\left(\frac{\beta}{2}\right) \\ \sqrt{3}\sin\left(\frac{\beta}{2}\right)\cos^2\left(\frac{\beta}{2}\right) & \frac{1}{2}(3\cos\beta-1)\cos\left(\frac{\beta}{2}\right) & -\frac{1}{2}(3\cos\beta+1)\sin\left(\frac{\beta}{2}\right) & \sqrt{3}\sin^2\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right) \\ \sqrt{3}\sin^2\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right) & \frac{1}{2}(3\cos\beta+1)\sin\left(\frac{\beta}{2}\right) & \frac{1}{2}(3\cos\beta-1)\cos\left(\frac{\beta}{2}\right) & -\sqrt{3}\sin\left(\frac{\beta}{2}\right)\cos^2\left(\frac{\beta}{2}\right) \\ \sin^3\left(\frac{\beta}{2}\right) & \sqrt{3}\sin^2\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right) & \sqrt{3}\sin\left(\frac{\beta}{2}\right)\cos^2\left(\frac{\beta}{2}\right) & \cos^3\left(\frac{\beta}{2}\right) \end{pmatrix}, \tag{7.444}$$

which can be reduced to

$$d^{3/2}(\beta) = \frac{\sin\beta}{2}\begin{pmatrix} \frac{\cos^2(\beta/2)}{\sin(\beta/2)} & -\sqrt{3}\cos\left(\frac{\beta}{2}\right) & \sqrt{3}\sin\left(\frac{\beta}{2}\right) & -\frac{\sin^2(\beta/2)}{\cos(\beta/2)} \\ \sqrt{3}\cos\left(\frac{\beta}{2}\right) & \frac{3\cos\beta-1}{2\sin(\beta/2)} & -\frac{3\cos\beta+1}{2\cos(\beta/2)} & \sqrt{3}\sin\left(\frac{\beta}{2}\right) \\ \sqrt{3}\sin\left(\frac{\beta}{2}\right) & \frac{3\cos\beta+1}{\cos(\beta/2)} & \frac{3\cos\beta-1}{2\sin(\beta/2)} & -\sqrt{3}\cos\left(\frac{\beta}{2}\right) \\ \frac{\sin^2(\beta/2)}{\cos(\beta/2)} & \sqrt{3}\sin\left(\frac{\beta}{2}\right) & \sqrt{3}\cos\left(\frac{\beta}{2}\right) & \frac{\cos^2(\beta/2)}{\sin(\beta/2)} \end{pmatrix}. \tag{7.445}$$

Following the method outlined in this problem, we can in principle find the matrix of any d-function. For instance, using the matrices of $d^{(1)}$ and $d^{(1/2)}$ along with the Clebsch–Gordan coefficients resulting from the addition of $j_1 = 1$ and $j_2 = 1$, we can find the matrix of $d^{(2)}(\beta)$.

Problem 7.11

Consider two nonidentical particles each with angular momenta 1 and whose Hamiltonian is given by

$$\hat{H} = \frac{\varepsilon_1}{\hbar^2}(\hat{\vec{L}}_1 + \hat{\vec{L}}_2)\cdot\hat{\vec{L}}_2 + \frac{\varepsilon_2}{\hbar^2}(\hat{L}_{1_z} + \hat{L}_{2_z})^2,$$

where ε_1 and ε_2 are constants having the dimensions of energy. Find the energy levels and their degeneracies for those states of the system whose total angular momentum is equal to $2\hbar$.

Solution

The total angular momentum of the system is obtained by coupling $l_1 = 1$ and $l_2 = 1$: $\hat{\vec{L}} = \hat{\vec{L}}_1 + \hat{\vec{L}}_2$. This leads to $\hat{\vec{L}}_1\cdot\hat{\vec{L}}_2 = \frac{1}{2}(\hat{L}^2 - \hat{L}_1^2 - \hat{L}_2^2)$, and when this is inserted into the system's Hamiltonian it yields

$$\hat{H} = \frac{\varepsilon_1}{\hbar^2}(\hat{\vec{L}}_1\cdot\hat{\vec{L}}_2 + \hat{L}_2^2) + \frac{\varepsilon_2}{\hbar^2}\hat{L}_z^2 = \frac{\varepsilon_1}{2\hbar^2}(\hat{L}^2 - \hat{L}_1^2 + \hat{L}_2^2) + \frac{\varepsilon_2}{\hbar^2}\hat{L}_z^2. \tag{7.446}$$

Notice that the operators $\hat{H}$, $\hat{L}_1^2$, $\hat{L}_2^2$, $\hat{L}^2$, and $\hat{L}_z$ mutually commute; we denote their joint eigenstates by $|\,l,\ m\rangle$. The energy levels of (7.446) are thus given by

$$E_{lm} = \frac{\varepsilon_1}{2}\Big[\,l(l+1) - l_1(l_1+1) + l_2(l_2+1)\Big] + \varepsilon_2 m^2 = \frac{\varepsilon_1}{2}l(l+1) + \varepsilon_2 m^2, \tag{7.447}$$

since $l_1 = l_2 = 1$.

The calculation of $|\, l,\ m\rangle$ in terms of the states $|l_1,\ m_1\rangle|l_2,\ m_2\rangle = |l_1, l_2;\ m_1, m_2\rangle$ was carried out in Problem 7.3, page 481; the states corresponding to a total angular momentum of $l = 2$ are given by

$$|\, 2,\ \pm 2\rangle = |1,1;\ \pm 1, \pm 1\rangle, \qquad |2,\ \pm 1\rangle = \frac{1}{\sqrt{2}}\Big(|1,1;\ \pm 1, 0\rangle + |1,1;\ 0, \pm 1\rangle\Big), \tag{7.448}$$

$$|\, 2,\ 0\rangle = \frac{1}{\sqrt{6}}\Big(|1,1;\ 1,-1\rangle + 2|1,1;\ 0,0\rangle + |1,1;\ -1,1\rangle\Big). \tag{7.449}$$

From (7.447) we see that the energy corresponding to $l = 2$ and $m = \pm 2$ is doubly degenerate, because the states $|\, 2,\ \pm 2\rangle$ have the same energy $E_{2,\pm 2} = 3\varepsilon_1 + 4\varepsilon_2$. The two states $|\, 2,\ \pm 1\rangle$ are also degenerate, for they correspond to the same energy $E_{2,\pm 1} = 3\varepsilon_1 + \varepsilon_2$. The energy corresponding to $|\, 2,\ 0\rangle$ is not degenerate: $E_{20} = 3\varepsilon_1$.

7.6 Exercises

Exercise 7.1
Show that the linear transformation $y = Rx$ where

$$R = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}, \qquad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

is a counterclockwise rotation of the Cartesian x_1x_2 coordinate system in the plane about the origin with an angle ϕ.

Exercise 7.2
Show that the nth power of the rotation matrix

$$R(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

is equal to

$$R^n(\phi) = \begin{pmatrix} \cos(n\phi) & -\sin(n\phi) \\ \sin(n\phi) & \cos(n\phi) \end{pmatrix}.$$

What is the geometrical meaning of this result?

Exercise 7.3
Show that the transformation of an operator $\widehat{\vec{R}}$ under the space displacement operator $\hat{U}(\widehat{\vec{A}}) = e^{i\,\widehat{\vec{A}}\cdot\widehat{\vec{P}}/\hbar}$, where $\widehat{\vec{P}}$ is the linear momentum operator, is given by $\widehat{\vec{R}}\,' = \widehat{\vec{R}} + \widehat{\vec{A}}$.

Exercise 7.4
The components A_j (with $j = x, y, z$) of a vector $\widehat{\vec{A}}$ transform under space rotations as $A'_i = R_{ij}A_j$, where R is the rotation matrix.

(a) Using the invariance of the scalar product of any two vectors (e.g., $\hat{\vec{A}} \cdot \hat{\vec{B}}$) under rotations, show that the rows and columns of the rotation matrix R are orthonormal to each other (i.e., show that $R_{lj}R_{lk} = \delta_{j,k}$).

(b) Show that the transpose of R is equal to the inverse of R and that the determinant of R is equal to ± 1.

Exercise 7.5

The operator corresponding to a rotation of angle θ about an axis $\vec{n}$ is given by

$$\hat{U}_{\vec{n}}(\theta) = e^{i\theta\, \vec{n}\cdot \hat{\vec{J}}/\hbar}.$$

Show that the matrix elements of the position operator $\hat{\vec{R}}$ are rotated through an infinitesimal rotation like $\vec{R}' = \hat{\vec{R}} + \theta\vec{n} \times \hat{\vec{R}}$.

Exercise 7.6

Consider the wave function of a particle $\psi(\vec{r}) = (\sqrt{2}x + \sqrt{2}y + z)f(r)$, where $f(r)$ is a spherically symmetric function.

(a) Is $\psi(\vec{r})$ an eigenfunction of $\hat{\vec{L}}^2$? If so, what is the eigenvalue?

(b) What are the probabilities for the particle to be found in the state $m_l = -1$, $m_l = 0$, and $m_l = 1$?

(c) If $\psi(\vec{r})$ is an energy eigenfunction with eigenvalues E and if $f(r) = 3r^2$, find the expression of the potential $V(r)$ to which this particle is subjected.

Exercise 7.7

Consider a particle whose wave function is given by

$$\psi(\vec{r}) = \left(\frac{1}{\sqrt{5}}Y_{11}(\theta,\varphi) - \frac{1}{5}Y_{1\text{-}1}(\theta,\varphi) + \frac{1}{\sqrt{2}}Y_{10}(\theta,\varphi)\right)f(r),$$

where $f(r)$ is a normalized radial function, i.e., $\int_0^\infty r^2 f^2(r)\,dr = 1$.

(a) Calculate the expectation values of $\hat{\vec{L}}^2$, $\hat{L}_z$, and $\hat{L}_x$ in this state.

(b) Calculate the expectation value of $V(\theta) = 2\cos^2\theta$ in this state.

(c) Find the probability that the particle will be found in the state $m_l = 0$.

Exercise 7.8

A particle of spin $\frac{1}{2}$ is in a d state of orbital angular momentum (i.e., $l = 2$). Work out the coupling of the spin and orbital angular momenta of this particle, and find all the states and the corresponding Clebsch–Gordan coefficients.

Exercise 7.9

The spin-dependent Hamiltonian of an electron–positron system in the presence of a uniform magnetic field in the z-direction ($\vec{B} = B\hat{k}$) can be written as

$$\hat{H} = \lambda \hat{\vec{S}}_1 \cdot \hat{\vec{S}}_2 + \left(\frac{eB}{mc}\right)\left(\hat{S}_{1_z} - \hat{S}_{2_z}\right),$$

where λ is a real number and $\vec{S}_1$ and $\vec{S}_2$ are the spin operators for the electron and the positron, respectively.

(a) If the spin function of the system is given by $|\frac{1}{2}, -\frac{1}{2}\rangle$, find the energy eigenvalues and their corresponding eigenvectors.

(b) Repeat (a) in the case where $\lambda = 0$, but $B \neq 0$.

(c) Repeat (a) in the case where $B = 0$, but $\lambda \neq 0$.

Exercise 7.10

(a) Show that $e^{-i\pi \hat{J}_z/2}\, e^{-i\pi \hat{J}_x}\, e^{i\pi \hat{J}_z/2} = e^{-i\pi \hat{J}_y}$.

(b) Prove $\hat{J}_- e^{-i\pi \hat{J}_x} = e^{-i\pi \hat{J}_x} \hat{J}_+$ and then show that $e^{-i\pi \hat{J}_x} \mid j,\ m\rangle = e^{-i\pi j} |j,\ -m\rangle$.

(c) Using (a) and (b), show that $e^{-i\pi \hat{J}_y} |j,\ m\rangle = (-1)^{j-m} |j,\ -m\rangle$.

Exercise 7.11

Using the commutation relations between the Pauli matrices, show that:

(a) $e^{i\alpha\sigma_y} \sigma_x e^{-i\alpha\sigma_y} = \sigma_x \cos(2\alpha) + \sigma_z \sin(2\alpha)$,

(b) $e^{i\alpha\sigma_z} \sigma_x e^{-i\alpha\sigma_z} = \sigma_x \cos(2\alpha) - \sigma_y \sin(2\alpha)$,

(c) $e^{i\alpha\sigma_x} \sigma_y e^{-i\alpha\sigma_x} = \sigma_y \cos(2\alpha) - \sigma_z \sin(2\alpha)$.

Exercise 7.12

(a) Show how $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$ transform under a rotation of (finite) angle α about the x-axis. Hint: take $\hat{R}_x(\alpha) = e^{i\alpha \hat{J}_x/\hbar}$.

(b) Using the results of part (a), determine how the angular momentum operator $\hat{\vec{J}}$ transforms under the rotation.

Exercise 7.13

(a) Show how the operator $\hat{J}_\pm$ transforms under a rotation of angle π about the x-axis.

(b) Use the result of part (a) to show that $\hat{J}_\pm e^{-i\pi \hat{J}_x/\hbar} = e^{-i\pi \hat{J}_x/\hbar} \hat{J}_\mp$.

Exercise 7.14

Consider a rotation of finite angle α about an axis $\vec{n}$ which transforms unit vector $\vec{a}$ into another unit vector $\vec{b}$. Show that $e^{-i\beta \hat{J}_b/\hbar} = e^{i\alpha \hat{J}_n/\hbar} e^{-i\beta \hat{J}_a/\hbar} e^{-i\alpha \hat{J}_n/\hbar}$.

Exercise 7.15

(a) Show that $e^{i\pi \hat{J}_y/2\hbar} \hat{J}_x e^{-i\pi \hat{J}_y/2\hbar} = \hat{J}_z$.

(b) Show also that $e^{i\pi \hat{J}_y/2\hbar} e^{i\alpha \hat{J}_z/\hbar} e^{-i\pi \hat{J}_y/2\hbar} = e^{i\alpha \hat{J}_z/\hbar}$.

(c) For any vector operator $\hat{\vec{A}}$, show that $e^{i\alpha \hat{J}_z/\hbar} \hat{A}_x e^{-i\alpha \hat{J}_z/\hbar} = \hat{A}_x \cos\alpha + \hat{A}_y \sin\alpha$.

Exercise 7.16

Using $\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2$ show that

$$d^{(j)}_{mm'}(\beta) = \sum_{m_1 m_2} \sum_{m_1' m_2'} \langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle \langle j_1, j_2;\ m_1' m_2' | j,\ m'\rangle d^{(j_1)}_{m_1 m_1'}(\beta) d^{(j_2)}_{m_2 m_2'}(\beta).$$

Exercise 7.17

Consider the tensor $A(\theta, \varphi) = \cos\theta \sin\theta \cos\varphi$.

(a) Calculate all the matrix elements $A_{m'm} = \langle l,\ m' \mid A \mid l,\ m\rangle$ for $l = 1$.

(b) Express $A(\theta, \varphi)$ in terms of the components of a spherical tensor of rank 2 (i.e., in terms of $Y_{2m}(\theta, \varphi)$).

(c) Calculate again all the matrix elements $A_{m'm}$, but this time using the Wigner–Eckart theorem. Compare these results with those obtained in (a). (The Clebsch–Gordan coefficients may be obtained from tables.)

Exercise 7.18

(a) Express xz/r^2 and $(x^2 - y^2)/r^2$ in terms of the components of a spherical tensor of rank 2.

(b) Using the Wigner–Eckart theorem, calculate the values of $\langle 1,\ 0 \mid xz/r^2 \mid 1,\ 1\rangle$ and $\langle 1,\ 1 \mid (x^2 - y^2)/r^2 \mid 1,\ -1\rangle$.

Exercise 7.19

Show that $\langle j,\ m' \mid e^{-i\beta \hat{J}_y/\hbar} \hat{J}_z^2 e^{i\beta \hat{J}_y/\hbar} \mid j,\ m'\rangle = \sum_{m=-j}^{m=j} m^2 \mid d^{(j)}_{m'm}(\beta) \mid^2$.

Exercise 7.20

Calculate the trace of the rotation matrix $D^{(1/2)}(\alpha, \beta, \gamma)$ for (a) $\beta = \pi$ and (b) $\alpha = \gamma = \pi$ and $\beta = 2\pi$.

Exercise 7.21

The quadrupole moment operator of a charge q is given by $\hat{Q}_{20} = q(3\hat{z}^2 - r^2)$. Write $\hat{Q}_{20}$ in terms of an irreducible spherical tensor of rank 2 and then express $\langle j,\ j|\hat{Q}_{20}|j,\ j\rangle$ in terms of j and the reduced matrix element $\langle j \parallel r^2 Y^{(2)} \parallel j\rangle$. *Hint:* You may use the coefficient $\langle j, 2;\ m, 0 \mid j,\ m\rangle = (-1)^{j-m}[3m^2 - j(j+1)]/\sqrt{(2j-1)j(j+1)(2j+3)}$.

Exercise 7.22

Prove the following commutation relations:

(a) $\left[J_x, [J_x, T^{(k)}_q]\right] = \sum_{q'} T^{(k)}_{q'} \langle k,\ q' \mid J_x^2 \mid k,\ q\rangle$,

(b) $\left[J_x, [J_x, T^{(k)}_q]\right] + \left[J_y, [J_y, T^{(k)}_q]\right] + \left[J_z, [J_z, T^{(k)}_q]\right] = k(k+1)\hbar^2 T^{(k)}_q$.

Exercise 7.23

Consider a spin $\frac{1}{2}$ particle which has an orbital angular momentum $l = 1$. Find all the Clebsch–Gordan coefficients involved in the addition of the orbital and spin angular momenta of this particle. *Hint:* The Clebsch–Gordan coefficient $\langle j_1, j_2;\ j_1, (j_2 - j_1) \mid j_2,\ j_2\rangle$ is real and positive.

Exercise 7.24

This problem deals with another derivation of the matrix elements of $d^{(1)}(\beta)$. Use the relation

$$d^{(j)}_{mm'}(\beta) = \sum_{m_1 m_2} \sum_{m_1' m_2'} \langle j_1, j_2;\ m_1, m_2 \mid j,\ m\rangle \langle j_1, j_2;\ m_1', m_2' \mid j,\ m'\rangle\ d^{(j_1)}_{m_1 m_1'}(\beta) d^{(j_2)}_{m_2 m_2'}(\beta)$$

for the case where $j_1 = j_2 = \frac{1}{2}$ along with the matrix elements of $d^{(1/2)}(\beta)$, which are given by (7.97), to derive all the matrix elements of $d^{(1)}(\beta)$.

Exercise 7.25

Consider the tensor $A(\theta, \varphi) = \sin^2\theta \cos(2\varphi)$.

(a) Calculate the reduced matrix element $\langle 2 \parallel Y_2 \parallel 2\rangle$. *Hint:* You may calculate explicitly $\langle 2,\ 1|Y_{20} \mid 2,\ 1\rangle$ and then use the Wigner–Eckart theorem to calculate it again.

(b) Express $A(\theta, \varphi)$ in terms of the components of a spherical tensor of rank 2 (i.e., in terms of $Y_{2m}(\theta, \varphi)$).

(c) Calculate $A_{m'\pm1} = \langle 2, m'|A|2, \pm1\rangle$ for $m' = \pm2, \pm1, 0$. You may need this Clebsch–Gordan coefficient: $\langle j, 2;\ m, 0 \mid j,\ m\rangle = [3m^2 - j(j+1)]/\sqrt{(2j-1)j(j+1)(2j+3)}$.

Exercise 7.26

(a) Calculate the reduced matrix element $\langle 1 \parallel Y_1 \parallel 2\rangle$. *Hint:* For this, you may need to calculate $\langle 1,\ 0 \mid Y_{10} \mid 2,\ 0\rangle$ directly and then from the Wigner–Eckart theorem.

(b) Using the Wigner–Eckart theorem and the relevant Clebsch–Gordan coefficients from the table, calculate $\langle 1,\ m|Y_{1m'}|2,\ m''\rangle$ for all possible values of m, m', and m''. *Hint:* You may find the integral $\int_0^\infty r^3 R_{21}^*(r) R_{32}(r)\, dr = \frac{64a_0}{15\sqrt{5}}\left(\frac{6}{5}\right)^5$ and the following coefficients useful:
$\langle j, 1;\ m, 0|(j-1),\ m\rangle = -\sqrt{(j-m)(j+m)/[j(2j+1)]}$,
$\langle j, 1;\ (m-1), 1|(j-1),\ m\rangle = \sqrt{(j-m)(j-m+1)/[2j(2j+1)]}$, and
$\langle j, 1;\ (m+1), -1|(j-1),\ m\rangle = \sqrt{(j+m)(j+m+1)/[2j(2j+1)]}$.

Exercise 7.27

A particle of spin $\frac{1}{2}$ is in a d state of orbital angular momentum (i.e., $l = 2$).

(a) What are its possible states of total angular momentum.

(b) If its Hamiltonian is given by $H = a + b\hat{\vec{L}} \cdot \hat{\vec{S}} + c\hat{\vec{L}}^2$, where a, b, and c are numbers, find the values of the energy for each of the different states of total angular momentum. Express your answer in terms of a, b, c.

Exercise 7.28

Consider an h-state electron. Calculate the Clebsch–Gordan coefficients involved in the following $\{|\ j,\ m\rangle\}$ states of the electron: $|\ \frac{11}{2},\ \frac{9}{2}\rangle, |\ \frac{11}{2},\ \frac{7}{2}\rangle, |\ \frac{9}{2},\ \frac{9}{2}\rangle, |\ \frac{9}{2},\ \frac{7}{2}\rangle$.

Exercise 7.29

Let the Hamiltonian of two nonidentical spin $\frac{1}{2}$ particles be

$$\hat{H} = \frac{\varepsilon_1}{\hbar^2}(\hat{\vec{S}}_1 + \hat{\vec{S}}_2) \cdot \hat{\vec{S}}_1 - \frac{\varepsilon_2}{\hbar}(\hat{S}_{1_z} + \hat{S}_{2_z}),$$

where ε_1 and ε_2 are constants having the dimensions of energy. Find the energy levels and their degeneracies.

Exercise 7.30

Find the energy levels and their degeneracies for a system of two nonidentical spin $\frac{1}{2}$ particles with Hamiltonian

$$\hat{H} = \frac{\varepsilon_0}{\hbar^2}(\hat{S}_1^2 + \hat{S}_2^2) - \frac{\varepsilon_0}{\hbar}(\hat{S}_{1_z} + \hat{S}_{2_z}),$$

where ε_0 is a constant having the dimensions of energy.

Exercise 7.31

Consider two nonidentical spin $s = \frac{1}{2}$ particles with Hamiltonian

$$\hat{H} = \frac{\varepsilon_0}{\hbar^2}(\hat{\vec{S}}_1 + \hat{\vec{S}}_2)^2 - \frac{\varepsilon_0}{\hbar^2}(\hat{S}_{1_z} + \hat{S}_{2_z})^2,$$

where ε_0 is a constant having the dimensions of energy. Find the energy levels and their degeneracies.

Exercise 7.32
Consider a system of three nonidentical particles, each of spin $s = \frac{1}{2}$, whose Hamiltonian is given by

$$\hat{H} = \frac{\varepsilon_1}{\hbar^2}(\hat{\vec{S}}_1 + \hat{\vec{S}}_3) \cdot \hat{\vec{S}}_2 + \frac{\varepsilon_2}{\hbar^2}(\hat{S}_{1_z} + \hat{S}_{2_z} + \hat{S}_{3_z})^2,$$

where ε_1 and ε_2 are constants having the dimensions of energy. Find the system's energy levels and their degeneracies.

Exercise 7.33
Consider a system of three nonidentical particles, each with angular momentum $\frac{3}{2}$. Find the possible values of the total spin S of this system and specify the number of angular momentum eigenstates corresponding to each value of S.

Chapter 8

Identical Particles

Up to this point, we have dealt mainly with the motion of a single particle. Now we want to examine how to describe systems with *many particles*. We shall focus on systems of *identical* particles and examine how to construct their wave functions.

8.1 Many-Particle Systems

Most physical systems—nucleons, nuclei, atoms, molecules, solids, fluids, gases, etc.—involve many particles. They are known as *many-particle* or *many-body* systems. While atomic, nuclear, and subnuclear systems involve intermediate numbers of particles ($\sim$ 2 to 300), solids, fluids, and gases are truly many-body systems, since they involve very large numbers of particles ($\sim 10^{23}$).

8.1.1 Schrödinger Equation

How does one describe the dynamics of a system of N particles? This description can be obtained from a generalization of the dynamics of a single particle. The state of a system of N spinless particles (we ignore their spin for the moment) is described by a wave function $\Psi(\vec{r}_1, \vec{r}_2, \ldots \vec{r}_N, t)$, where $|\Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N, t)|^2 d^3r_1\, d^3r_2 \cdots d^3r_N$ represents the probability at time t of finding particle 1 in the volume element d^3r_1 centered about $\vec{r}_1$, particle 2 in the volume d^3r_2 about $\vec{r}_2, \ldots$, and particle N in the volume d^3r_N about $\vec{r}_N$. The normalization condition of the state is given by

$$\int d^3r_1 \int d^3r_2 \cdots \int |\Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N, t)|^2 d^3r_N = 1. \tag{8.1}$$

The wave function Ψ evolves in time according to the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N, t) = \hat{H} \Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N, t). \tag{8.2}$$

The form of $\hat{H}$ is obtained by generalizing the one-particle Hamiltonian $\vec{P}^2/(2m) + \vec{V}(\vec{r})$ to N

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

particles:

$$\hat{H} = \sum_{j=1}^{N} \frac{\hat{\vec{p}}_j^2}{2m_j} + \hat{V}(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N, t) = -\sum_{j=1}^{N} \frac{\hbar^2}{2m_j}\nabla_j^2 + \hat{V}(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N, t), \tag{8.3}$$

where m_j and $\hat{\vec{p}}_j$ are the mass and the momentum of the jth particle and $\hat{V}$ is the operator corresponding to the total potential energy ($\hat{V}$ accounts, on the one hand, for all forms of interactions (internal and external), the mutual interactions between the various particles of the system, and, on the other hand, for the interactions of the particles with the outside world).

The formalism of quantum mechanics for an N-particle system can be, in principle, inferred from that of a single particle. Operators corresponding to different particles commute; for instance, the commutation relations between the position and momentum operators are

$$[\hat{X}_j, \hat{P}_{x_k}] = i\hbar\delta_{j,k}, \qquad [\hat{X}_j, \hat{X}_k] = [\hat{P}_{x_j}, \hat{P}_{x_k}] = 0 \qquad (j,k = 1,2,3,\ldots,N), \tag{8.4}$$

where $\hat{X}_j$ is the x-position operator of the jth particle, and $\hat{P}_{x_k}$ the x-momentum operator of the kth particle; similar relations can be obtained for the y and z components.

Stationary states

In the case where the potential $\hat{V}$ is *time independent*, the solutions of (8.2) are given by *stationary* states

$$\Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N, t) = \psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)\, e^{-iEt/\hbar}, \tag{8.5}$$

where E is the total energy of the system and ψ is the solution to the time-independent Schrödinger equation $\hat{H}\psi = E\psi$, i.e.,

$$\left[-\sum_{j=1}^{N} \frac{\hbar^2}{2m_j}\vec{\nabla}_j^2 + V(\vec{r}_1, \ldots, \vec{r}_N)\right]\psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) = E\,\psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N). \tag{8.6}$$

The properties of stationary states for a single particle also apply to N-particle systems. For instance, the probability density $\langle\psi \mid \psi\rangle$, the probability current density $\vec{j}$, and the expectation values of time-independent operators are conserved, since they do not depend on time:

$$\langle\Psi \mid \hat{A} \mid \Psi\rangle = \langle\psi \mid \hat{A} \mid \psi\rangle = \int d^3r_1 \int d^3r_2 \cdots \int \psi^*(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)\hat{A}\psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)\, d^3r_N. \tag{8.7}$$

In particular, the energy of a stationary state is conserved.

Multielectron atoms

As an illustration, let us consider an atom with Z electrons. If $\vec{R}$ is used to represent the position of the center of mass of the nucleus, the wave function of the atom depends on $3(Z+1)$ coordinates $\psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_Z, \vec{R})$, where $\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_Z$ are the position vectors of the Z electrons. The time-independent Schrödinger equation for this atom, neglecting contributions from the spin–orbit correction, the relativistic correction, and similar terms, is given by

$$\left[-\frac{\hbar^2}{2m_e}\sum_{i=1}^{Z}\vec{\nabla}_{r_i}^2 - \frac{\hbar^2}{2M}\vec{\nabla}_R^2 - \sum_{i=1}^{Z}\frac{Ze^2}{|\vec{r}_i - \vec{R}|} + \sum_{i>j}\frac{e^2}{|\vec{r}_i - \vec{r}_j|}\right]\psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_Z, \vec{R})$$
$$= E\psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_Z, \vec{R}), \tag{8.8}$$

where M is the mass of the nucleus and $-\hbar^2\vec{\nabla}_R^2/2M$ is its kinetic energy operator. The term $-\sum_{i=1}^{Z} Ze^2/|\vec{r}_i - \vec{R}|$ represents the attractive Coulomb interaction of each electron with the nucleus and $\sum_{i>j} e^2/|\vec{r}_i - \vec{r}_j|$ is the repulsive Coulomb interaction between the ith and the jth electrons; $|\vec{r}_i - \vec{r}_j|$ is the distance separating them. As these (Coulomb) interactions are independent of time, the states of atoms are stationary.

We should note that the Schrödinger equations (8.3), (8.6), and (8.8) are all many-particle differential equations. As these equations cannot be separated into one-body equations, it is difficult, if not impossible, to solve them. For the important case where the N particles of the system do not interact—this is referred to as an *independent particle system*—the Schrödinger equation can be trivially reduced to N one-particle equations (Section 8.1.3); we have seen how to solve these equations exactly (Chapters 4 and 6) and approximately (Chapters 9 and 10).

8.1.2 Interchange Symmetry

Although the exact eigenstates of the many-body Hamiltonian (8.3) are generally impossible to obtain, we can still infer some of their properties by means of symmetry schemes. Let ξ_i represent the coordinates (position $\vec{r}_i$, spin $\vec{s}_i$, and any other internal degrees of freedom such as isospin, color, flavor) of the ith particle and let $\psi(\xi_1, \xi_2, \ldots, \xi_N)$ designate the wave function of the N-particle system.

We define a *permutation operator* (also called *exchange operator*) $\hat{P}_{ij}$ as an operator that, when acting on an N-particle wave function $\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N)$, interchanges the ith and the jth particles

$$\hat{P}_{ij}\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots \xi_N) = \psi(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N); \tag{8.9}$$

i and j are arbitrary ($i, j = 1, 2, \cdots, N$). Since

$$\begin{aligned} \hat{P}_{ji}\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots \xi_N) &= \psi(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N) \\ &= \hat{P}_{ij}\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots \xi_N), \end{aligned} \tag{8.10}$$

we have $\hat{P}_{ij} = \hat{P}_{ji}$. In general, permutation operators do not commute:

$$\hat{P}_{ij}\hat{P}_{kl} \neq \hat{P}_{kl}\hat{P}_{ij} \quad \text{or} \quad [\hat{P}_{ij}, \hat{P}_{kl}] \neq 0 \quad (ij \neq kl). \tag{8.11}$$

For instance, in the case of a four-particle state $\psi(\xi_1, \xi_2, \xi_3, \xi_4) = (3\xi_4/\xi_2\xi_3)e^{-i\xi_1}$, we have

$$\hat{P}_{12}\hat{P}_{14}\psi(\xi_1, \xi_2, \xi_3, \xi_4) = \hat{P}_{12}\psi(\xi_4, \xi_2, \xi_3, \xi_1) = \psi(\xi_2, \xi_4, \xi_3, \xi_1) = \frac{3\xi_1}{\xi_4\xi_3}e^{-i\xi_2}, \tag{8.12}$$

$$\hat{P}_{14}\hat{P}_{12}\psi(\xi_1, \xi_2, \xi_3, \xi_4) = \hat{P}_{14}\psi(\xi_2, \xi_1, \xi_3, \xi_4) = \psi(\xi_4, \xi_1, \xi_3, \xi_2) = \frac{3\xi_2}{\xi_1\xi_3}e^{-i\xi_4}. \tag{8.13}$$

Since two successive applications of $\hat{P}_{ij}$ leave the wave function unchanged,

$$\begin{aligned} \hat{P}_{ij}^2\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) &= \hat{P}_{ij}\psi(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N) \\ &= \psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N), \end{aligned} \tag{8.14}$$

we have $\hat{P}_{ij}^2 = 1$; hence $\hat{P}_{ij}$ has two eigenvalues ± 1:

$$\hat{P}_{ij}\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) = \pm\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots \xi_N). \tag{8.15}$$

The wave functions corresponding to the eigenvalue +1 are *symmetric* and those corresponding to −1 are *antisymmetric* with respect to the interchange of the pair (i, j). Denoting these functions by ψ_s and ψ_a, respectively, we have

$$\psi_s(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) = \psi_s(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N), \tag{8.16}$$

$$\psi_a(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) = -\psi_a(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N). \tag{8.17}$$

Example 8.1
Specify the symmetry of the following functions:

(a) $\psi(x_1, x_2) = 4(x_1 - x_2)^2 + \frac{10}{x_1^2 + x_2^2}$,

(b) $\phi(x_1, x_2) = -\frac{3(x_1 - x_2)}{2(x_1 - x_2)^2 + 7}$,

(c) $\chi(x_1, x_2, x_3) = 6x_1x_2x_3 + \frac{x_1^2 + x_2^2 + x_3^2 - 1}{2x_1^3 + 2x_2^3 + 2x_3^3 + 5}$,

(d) $\Phi(x_1, x_2) = \frac{1}{x_2 + 3}e^{-|x_1|}$.

Solution

(a) The function $\psi(x_1, x_2)$ is symmetric, since $\psi(x_2, x_1) = \psi(x_1, x_2)$.

(b) The function $\phi(x_1, x_2)$ is antisymmetric, since $\phi(x_2, x_1) = -\phi(x_1, x_2)$, and ϕ is zero when $x_1 = x_2$: $\phi(x_1, x_1) = 0$.

(c) The function $\chi(x_1, x_2, x_3)$ is symmetric because

$$\begin{aligned}\chi(x_1, x_2, x_3) &= \chi(x_1, x_3, x_2) = \chi(x_2, x_1, x_3) = \chi(x_2, x_3, x_1) \\ &= \chi(x_3, x_1, x_2) = \chi(x_3, x_2, x_1).\end{aligned} \tag{8.18}$$

(d) The function $\Phi(x_2, x_1)$ is neither symmetric nor antisymmetric, since $\Phi(x_2, x_1) = \frac{1}{x_1 + 3}e^{-|x_2|} \neq \pm\Phi(x_1, x_2)$.

8.1.3 Systems of Distinguishable Noninteracting Particles

For a system of N noninteracting particles that are distinguishable—each particle has a different mass m_i and experiences a different potential $\hat{V}_i(\xi_i)$—the potential $\hat{V}$ is given by

$$\hat{V}(\xi_1, \xi_2, \ldots, \xi_N) = \sum_{i=1}^{N} \hat{V}_i(\xi_i) \tag{8.19}$$

and the Hamiltonian of this system of N independent particles by

$$\hat{H} = \sum_{i=1}^{N} \hat{H}_i = \sum_{i=1}^{N} \left[-\frac{\hbar^2}{2m_i}\nabla_i^2 + \hat{V}_i(\xi_i) \right], \tag{8.20}$$

where $\hat{H}_i = -\hbar^2\nabla_i^2/2m_i + \hat{V}_i(\xi_i)$ is the Hamiltonian of the ith particle, known as the single particle Hamiltonian. The Hamiltonians of different particles commute $[\hat{H}_i, \hat{H}_j] = 0$, since $[\hat{X}_i, \hat{X}_j] = [\hat{P}_i, \hat{P}_j] = 0$.

The Schrödinger equation of the N-particle system

$$\hat{H}\psi_{n_1,n_2,\cdots,n_N}(\xi_1, \xi_2, \cdots, \xi_N) = E_{n_1,n_2,\cdots,n_N}\psi_{n_1,n_2,\cdots,n_N}(\xi_1, \xi_2, \cdots, \xi_N) \tag{8.21}$$

separates into N one-particle equations

$$\left[-\frac{\hbar^2}{2m_i}\nabla_i^2 + \hat{V}_i(\xi_i)\right]\psi_{n_i}(\xi_i) = \varepsilon_{n_i}\psi_{n_i}(\xi_i), \tag{8.22}$$

with

$$E_{n_1,n_2,\cdots,n_N} = \varepsilon_{n_1} + \varepsilon_{n_2} + \cdots \varepsilon_{n_N} = \sum_{i=1}^{N} \varepsilon_{n_i} \tag{8.23}$$

and

$$\psi_{n_1,n_2,\cdots,n_N}(\xi_1, \xi_2, \cdots, \xi_N) = \psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\cdots\psi_{n_N}(\xi_N) = \prod_{i=1}^{N} \psi_{n_i}(\xi_i). \tag{8.24}$$

We see that, when the interactions are neglected, the N-particle Schrödinger equation separates into N one-particle Schrödinger equations. The solutions of these equations yield the single-particle energies ε_{n_i} and states $\psi_{n_i}(\xi_i)$; the single-particle states are also known as the *orbitals*. The total energy is the sum of the single-particle energies and the total wave function is the product of the orbitals. The number n_i designates the set of all quantum numbers of the ith particle. Obviously, each particle requires one, two, or three quantum numbers for its full description, depending on whether the particles are moving in a one-, two-, or three-dimensional space; if the spin were considered, we would need to add another quantum number. For instance, if the particles moved in a one-dimensional harmonic oscillator, n_i would designate the occupation number of the ith particle. But if the particles were the electrons of an atom, then n_i would stand for four quantum numbers: the radial, orbital, magnetic, and spin quantum numbers $N_i l_i m_{l_i} m_{s_i}$.

Example 8.2

Find the energy levels and wave functions of a system of four distinguishable spinless particles placed in an infinite potential well of size a. Use this result to infer the energy and the wave function of the ground state and the first excited state.

Solution

Each particle moves in a potential which is defined by $\hat{V}_i(x_i) = 0$ for $0 \le x_i \le a$ and $\hat{V}_i(x_i) = \infty$ for the other values of x_i. In this case the Schrödinger equation of the four-particle system:

$$\sum_{i=1}^{4}\left[-\frac{\hbar^2}{2m_i}\frac{d^2}{dx_i^2}\right]\psi_{n_1,n_2,n_3,n_4}(x_1, x_2, x_3, x_4) = E_{n_1,n_2,n_3,n_4}\psi_{n_1,n_2,n_3,n_4}(x_1, x_2, x_3, x_4), \tag{8.25}$$

separates into four one-particle equations

$$-\frac{\hbar^2}{2m_i}\frac{d^2\psi_{n_i}(x_i)}{dx_i^2} = \varepsilon_{n_i}\psi_{n_i}(x_i), \qquad i = 1, 2, 3, 4, \tag{8.26}$$

with

$$\varepsilon_{n_i} = \frac{\hbar^2\pi^2 n_i^2}{2m_i a^2}, \qquad \psi_{n_i}(x_i) = \sqrt{\frac{2}{a}}\sin\left(\frac{n_i\pi}{a}x_i\right). \tag{8.27}$$

The total energy and wave function are given by

$$E_{n_1,n_2,n_3,n_4} = \frac{\hbar^2\pi^2}{2a^2}\left(\frac{n_1^2}{m_1} + \frac{n_2^2}{m_2} + \frac{n_3^2}{m_3} + \frac{n_4^2}{m_4}\right), \tag{8.28}$$

$$\psi_{n_1,n_2,n_3,n_4}(x_1,x_2,x_3,x_4) = \frac{4}{a^2}\sin\left(\frac{n_1\pi}{a}x_1\right)\sin\left(\frac{n_2\pi}{a}x_2\right)\sin\left(\frac{n_3\pi}{a}x_3\right)\sin\left(\frac{n_4\pi}{a}x_4\right). \tag{8.29}$$

The ground state corresponds to the case where all four particles occupy their respective ground state orbitals, $n_1 = n_2 = n_3 = n_4 = 1$. The ground state energy and wave function are thus given by

$$E_{1,1,1,1} = \frac{\hbar^2\pi^2}{2a^2}\left(\frac{1}{m_1} + \frac{1}{m_2} + \frac{1}{m_3} + \frac{1}{m_4}\right), \tag{8.30}$$

$$\psi_{1,1,1,1}(x_1,x_2,x_3,x_4) = \frac{4}{a^2}\sin\left(\frac{\pi}{a}x_1\right)\sin\left(\frac{\pi}{a}x_2\right)\sin\left(\frac{\pi}{a}x_3\right)\sin\left(\frac{\pi}{a}x_4\right). \tag{8.31}$$

The first excited state is somewhat tricky. Since it corresponds to the next higher energy level of the system, it must correspond to the case where the particle having the largest mass occupies its first excited state while the other three particles remain in their respective ground states. For argument's sake, if the third particle were the most massive, the first excited state would correspond to the configuration $n_1 = n_2 = n_4 = 1$ and $n_3 = 2$; the energy and wave function of the first excited state would then be given by

$$E_{1,1,2,1} = \frac{\hbar^2\pi^2}{2a^2}\left(\frac{1}{m_1} + \frac{1}{m_2} + \frac{4}{m_3} + \frac{1}{m_4}\right), \tag{8.32}$$

$$\psi_{1,1,2,1}(x_1,x_2,x_3,x_4) = \frac{4}{a^2}\sin\left(\frac{\pi}{a}x_1\right)\sin\left(\frac{\pi}{a}x_2\right)\sin\left(\frac{2\pi}{a}x_3\right)\sin\left(\frac{\pi}{a}x_4\right). \tag{8.33}$$

Continuing in this way, we can obtain the entire energy spectrum of this system.

8.2 Systems of Identical Particles

8.2.1 Identical Particles in Classical and Quantum Mechanics

In classical mechanics, when a system is made of identical particles, it is possible to identify and distinguish each particle from the others. That is, although all particles have the same physical properties, we can "tag" each classical particle and follow its motion along a path. For instance, each particle can be colored differently from the rest; hence we can follow the trajectory of each particle separately at each time. *Identical classical particles, therefore, do not lose their identity; they are distinguishable.*

In quantum mechanics, however, identical particles are truly indistinguishable. The underlying basis for this is twofold. First, to describe a particle, we cannot specify more than a

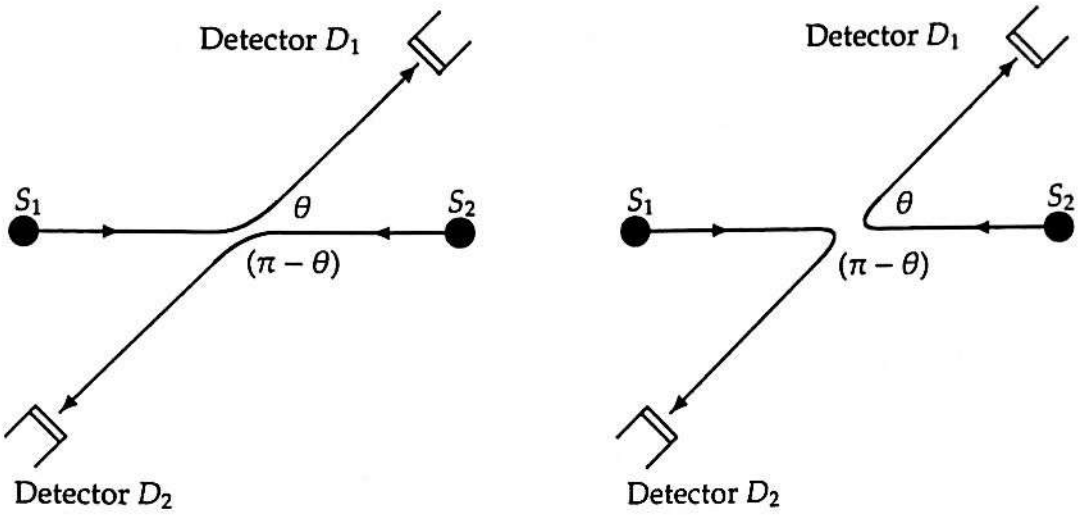


Figure 8.1: When scattering two *identical* particles in the center of mass frame, it is impossible to forcast with certitude whether the particles scatter according to the first process or to the second. For instance, we cannot tell whether the particle fired from source S_1 will make it to detector D_1 or to D_2.

complete set of commuting observables. In particular, there exists no mechanism to tag the particles as in classical mechanics. Second, due to the uncertainty principle, the concept of the path of a particle becomes meaningless. Even if the position of a particle is exactly determined at a given time, it is not possible to specify its coordinates at the next instant. Thus, *identical particles lose their identity (individuality) in quantum mechanics.*

To illustrate this, consider an experiment in which we scatter two identical particles. As displayed in Figure 8.1, after particles 1 and 2 (fired from the sources S_1 and S_2) have scattered, it is impossible to distinguish between the first and the second outcomes. That is, we cannot determine experimentally the identity of the particles that are collected by each detector. For instance, we can in no way tell whether it is particle 1 or particle 2 that has reached detector D_1. We can only say that a particle has reached detector D_1 and another has made it to D_2, but have no information on their respective identities. There exists no experimental mechanism that allows us to follow the motion of each particle from the time it is fired out of the source till it reaches the detector. This experiment shows that *the individuality of a microscopic particle is lost the moment it is mixed with other similar particles.*

Having discussed the indistinguishability concept on a two-particle system, let us now study this concept on larger systems. For this, consider a system of N *identical* particles whose wave function is $\psi(\xi_1, \xi_2, \ldots, \xi_N)$.

The moment these N particles are mixed together, no experiment can determine which particle has the coordinates ξ_1, or which one has ξ_2, and so on. It is impossible to specify experimentally the identity of the particle which is located at ξ_1, or that located at ξ_2, and so on. The only measurements we can perform are those that specify the probability for a certain particle to be located at ξ_1, another at ξ_2, and so on, but we can never make a distinction as to which particle is which.

As a result, the probability must remain unchanged by an interchange of the particles. For

instance, an interchange of particles i and j will leave the probability density unaffected:

$$\left|\psi(\xi_1, \xi_2, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N)\right|^2 = \left|\psi(\xi_1, \xi_2, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N)\right|^2; \tag{8.34}$$

hence we have

$$\psi(\xi_1, \xi_2, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) = \pm\psi(\xi_1, \xi_2, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N). \tag{8.35}$$

This means that the wave function of a system of N identical particles is either symmetric or antisymmetric under the interchange of a pair of particles. We will deal with the implications of this result in Section 8.2.3. We will see that the sign in (8.35) is related to the spin of the particles: the negative sign corresponds to particles with half-odd-integral spin and the positive sign corresponds to particles with integral spin; that is, the wave functions of particles with integral spins are symmetric and the wave functions of particles with half-odd-integral spins are antisymmetric. In fact, experimental observations show that, in nature, particles come in two classes:

- Particles with *integral spin*, $S_i = 0, 1\hbar, 2\hbar, 3\hbar, \ldots$, such as photons, pions, alpha particles. These particles are called *bosons*.
- Particles with *half-odd-integral spin*, $S_i = \hbar/2, 3\hbar/2, 5\hbar/2, 7\hbar/2, \ldots$, such as quarks, electrons, positrons, protons, neutrons. These particles are called *fermions*.

That is, particles occurring in nature are either bosons or fermions.

Before elaborating more on the properties of bosons and fermions, let us present a brief outline on the interchange (permutation) symmetry.

8.2.2 Exchange Degeneracy

How does the interchange symmetry affect operators such as the Hamiltonian? Since the Coulomb potential, which results from electron–electron and electron–nucleus interactions,

$$V(\vec{r}_1, \vec{r}_2, \vec{r}_3, \ldots, \vec{r}_Z) = -\sum_{i=1}^{Z} \frac{Z e^2}{|\vec{r}_i - \vec{R}|} + \sum_{i>j} \frac{e^2}{|\vec{r}_i - \vec{r}_j|}, \tag{8.36}$$

is invariant under the permutation of any pair of electrons, the Hamiltonian (8.8) is also invariant under such permutations. This symmetry also applies to the orbital, spin, and angular momenta of an atom. We may thus use this symmetry to introduce another definition of the *identicalness* of particles. *The N particles of a system are said to be identical if the various observables of the system* (such as the Hamiltonian $\hat{H}$, the angular momenta, and so on) *are symmetrical when any two particles are interchanged.* If these operators were not symmetric under particle interchange, the particles would be distinguishable.

The invariance of the Hamiltonian under particle interchanges is not without physical implications: the eigenvalues of $\hat{H}$ are *degenerate*. The wave functions corresponding to all possible electron permutations have the same energy E: $\hat{H}\psi = E\psi$. This is known as the *exchange degeneracy*. For instance, the degeneracy associated with a system of two identical particles is equal to 2, since $\psi(\xi_1, \xi_2)$ and $\psi(\xi_2, \xi_1)$ correspond to the same energy E.

So the Hamiltonian of a system of N identical particles ($m_i = m$) is completely symmetric with respect to the coordinates of the particles:

$$\begin{aligned}\hat{H}(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) &= \sum_{k=1}^{N} \frac{\vec{p}_k^2}{2m} + \hat{V}(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) \\ &= \hat{H}(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N),\end{aligned} \tag{8.37}$$

because $\hat{V}$ is invariant under the permutation of any pair of particles $i \longleftrightarrow j$:

$$\hat{V}(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) = \hat{V}(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N). \tag{8.38}$$

This property can also be ascertained by showing that $\hat{H}$ commutes with the particle interchange operator $\hat{P}_{ij}$. If ψ is eigenstate to $\hat{H}$ with eigenvalue E, we can write

$$\begin{aligned}\hat{H}\hat{P}_{ij}\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) &= \hat{H}\psi(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N) \\ &= E\psi(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N) = E\hat{P}_{ij}\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) \\ &= \hat{P}_{ij}E\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) = \hat{P}_{ij}\hat{H}\psi(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N)\end{aligned} \tag{8.39}$$

or

$$[\hat{H}, \hat{P}_{ij}] = 0. \tag{8.40}$$

Therefore, $\hat{P}_{ij}$ is a *constant of the motion*. That is, if we start with a wave function that is symmetric (antisymmetric), it will remain so for all subsequent times. Moreover, since $\hat{P}_{ij}$ and $\hat{H}$ commute, they possess a complete set of functions that are joint eigenstates of both. As shown in (8.15) to (8.17), these eigenstates have definite parity, either symmetric or antisymmetric.

8.2.3 Symmetrization Postulate

We have shown in (8.35) that the wave function of a system of N identical particles is either symmetric or antisymmetric under the interchange of any pair of particles:

$$\psi(\xi_1, \xi_2, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_N) = \pm\psi(\xi_1, \xi_2, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_N). \tag{8.41}$$

This result, which turns out to be supported by experimental evidence, is the very essence of the *symmetrization postulate* which stipulates that, in nature, the states of systems containing N identical particles are *either totally symmetric or totally antisymmetric* under the interchange of any pair of particles and that states with mixed symmetry do not exist. Besides that, this postulate states two more things:

- Particles with integral spins, or bosons, have symmetric states.
- Particles with half-odd-integral spins, or fermions, have antisymmetric states.

Fermions are said to obey *Fermi–Dirac statistics* and bosons to obey *Bose–Einstein statistics*. So the wave function of a system of identical bosons is totally symmetric and the wave function of a system of identical fermions is totally antisymmetric.

Composite particles

The foregoing discussion pertains to identical particles that are "simple" or elementary such as quarks, electrons, positrons, muons, and so on. Let us now discuss the symmetry of systems of identical *composite* "particles" where each particle is composed of two or more identical elementary particles. For instance, alpha particles, which consist of nuclei that are composed of two neutrons and two protons each, are a typical example of composite particles. A system of N hydrogen atoms can also be viewed as a system of identical composite particles where each "particle" (atom) consists of a proton and an electron. Protons, neutrons, pions, etc., are themselves composite particles, because protons and neutrons consist of three quarks and pions consist of two. *Quarks* are elementary spin $\frac{1}{2}$ particles.

Composite particles have spin. The spin of each composite particle can be obtained by adding the spins of its constituents. If the total spin of the composite particle is half-odd-integer, this particle behaves like a fermion, and hence it obeys Fermi–Dirac statistics. If, on the other hand, its resultant spin is integer, it behaves like a boson and obeys Bose–Einstein statistics. In general, if the composite particle has an odd number of fermions; it is then a fermion, otherwise it is a boson. For instance, nucleons are fermions because they consist of three quarks; mesons are bosons because they consist of two quarks. For another illustrative example, let us consider the isotopes ^{4}He and ^{3}He of the helium atom: ^{4}He, which is called an alpha particle, is a boson for it consists of four nucleons (two protons and two neutrons), while ^{3}He is a fermion since it consists of three nucleons (one neutron and two protons). The hydrogen atom consists of two fermions (an electron and a proton), so it is a boson.

8.2.4 Constructing Symmetric and Antisymmetric Functions

Since the wave functions of systems of identical particles are either totally symmetric or totally antisymmetric, it is appropriate to study the formalism of how to construct wave functions that are totally symmetric or totally antisymmetric starting from asymmetric functions. For simplicity, consider first a system of two identical particles. Starting from any normalized asymmetric wave function $\psi(\xi_1, \xi_2)$, we can construct symmetric wave functions $\psi_s(\xi_1, \xi_2)$ as

$$\psi_s(\xi_1, \xi_2) = \frac{1}{\sqrt{2}}\Big[\psi(\xi_1, \xi_2) + \psi(\xi_2, \xi_1)\Big] \tag{8.42}$$

and antisymmetric wave functions $\psi_a(\xi_1, \xi_2)$ as

$$\psi_a(\xi_1, \xi_2) = \frac{1}{\sqrt{2}}\Big[\psi(\xi_1, \xi_2) - \psi(\xi_2, \xi_1)\Big], \tag{8.43}$$

where $1/\sqrt{2}$ is a normalization factor.

Similarly, for a system of three identical particles, we can construct ψ_s and ψ_a from an

asymmetric function ψ as follows:

$$\psi_s(\xi_1,\xi_2,\xi_3) = \frac{1}{\sqrt{6}}\Big[\psi(\xi_1,\xi_2,\xi_3)+\psi(\xi_1,\xi_3,\xi_2)+\psi(\xi_2,\xi_3,\xi_1) + \psi(\xi_2,\xi_1,\xi_3)+\psi(\xi_3,\xi_1,\xi_2)+\psi(\xi_3,\xi_2,\xi_1)\Big], \tag{8.44}$$

$$\psi_a(\xi_1,\xi_2,\xi_3) = \frac{1}{\sqrt{6}}\Big[\psi(\xi_1,\xi_2,\xi_3)-\psi(\xi_1,\xi_3,\xi_2)+\psi(\xi_2,\xi_3,\xi_1) - \psi(\xi_2,\xi_1,\xi_3)+\psi(\xi_3,\xi_1,\xi_2)-\psi(\xi_3,\xi_2,\xi_1)\Big]. \tag{8.45}$$

Continuing in this way, we can in principle construct symmetric and antisymmetric wave functions for any system of N identical particles.

8.2.5 Systems of Identical Noninteracting Particles

In the case of a system of N noninteracting identical particles, where all particles have equal mass $m_i = m$ and experience the same potential $\hat{V}_i(\xi_i) = \hat{V}(\xi_i)$, the Schrödinger equation of the system separates into N identical one-particle equations

$$\left[-\frac{\hbar^2}{2m}\nabla_i^2 + \hat{V}(\xi_i)\right]\psi_{n_i}(\xi_i) = \varepsilon_{n_i}\psi_{n_i}(\xi_i). \tag{8.46}$$

Whereas the energy is given, like the case of a system of N distinguishable particles, by a sum of the single-particle energies $E_{n_1,n_2,\ldots,n_N} = \sum_{i=1}^{N}\varepsilon_{n_i}$, the wave function can no longer be given by a simple product $\psi_{n_1,n_2,\ldots,n_N}(\xi_1,\xi_2,\ldots,\xi_N) = \prod_{i=1}^{N}\psi_{n_i}(\xi_i)$ for at least two reasons. First, if the wave function is given by such a product, it would imply that particle 1 is in the state ψ_{n_1}, particle 2 in the state ψ_{n_2}, ..., and particle N in the state ψ_{n_N}. This, of course, makes no sense since all we know is that one of the particles is in the state ψ_{n_1}, another in ψ_{n_2}, and so on; since the particles are identical, there is no way to tell which particle is in which state. If, however, the particles were distinguishable, then their total wave function would be given by such a product, as shown in (8.24). The second reason why the wave function of a system of identical particles cannot be given by $\prod_{i=1}^{N}\psi_{n_i}(\xi_i)$ has to do with the fact that such a product has, in general, no *definite* symmetry—a mandatory requirement for systems of N identical particles whose wave functions are either symmetric or antisymmetric. We can, however, extend the method of Section 8.2.4 to construct totally symmetric and totally antisymmetric wave functions from the single-particle states $\psi_{n_i}(\xi_i)$. For this, we are going to show how to construct symmetrized and antisymmetrized wave functions for systems of two, three, and N noninteracting identical particles.

8.2.5.1 Wave Function of Two-Particle Systems

By analogy with (8.42) and (8.43), we can construct the symmetric and antisymmetric wave functions for a system of two identical, noninteracting particles in terms of the single-particle

wave functions as follows:

$$\psi_s(\xi_1,\xi_2) = \frac{1}{\sqrt{2}}\Big[\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)+\psi_{n_1}(\xi_2)\psi_{n_2}(\xi_1)\Big], \tag{8.47}$$

$$\psi_a(\xi_1,\xi_2) = \frac{1}{\sqrt{2}}\Big[\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)-\psi_{n_1}(\xi_2)\psi_{n_2}(\xi_1)\Big], \tag{8.48}$$

where we have supposed that $n_1 \neq n_2$. When $n_1 = n_2 = n$ the symmetric wave function is given by $\psi_s(\xi_1,\xi_2) = \psi_n(\xi_1)\psi_n(\xi_2)$ and the antisymmetric wave function is zero; we will deal later with the reason why $\psi_a(\xi_1,\xi_2) = 0$ whenever $n_1 = n_2$.

Note that we can rewrite ψ_s as

$$\boxed{\psi_s(\xi_1,\xi_2) = \frac{1}{\sqrt{2!}}\sum_P \hat{P}\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2),} \tag{8.49}$$

where $\hat{P}$ is the permutation operator and where the sum is over all possible permutations (here we have only two possible ones). Similarly, we can write ψ_a as

$$\psi_a(\xi_1,\xi_2) = \frac{1}{\sqrt{2!}}\sum_P (-1)^P \hat{P}\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2), \tag{8.50}$$

where $(-1)^P$ is equal to +1 for an even permutation (i.e., when we interchange both ξ_1 and ξ_2 and also n_1 and n_2) and equal to −1 for an odd permutation (i.e., when we permute ξ_1 and ξ_2 but not n_1, n_2, and vice versa). Note that we can rewrite ψ_a of (8.48) in the form of a determinant

$$\boxed{\psi_a(\xi_1,\xi_2) = \frac{1}{\sqrt{2!}}\begin{vmatrix}\psi_{n_1}(\xi_1) & \psi_{n_1}(\xi_2)\\ \psi_{n_2}(\xi_1) & \psi_{n_2}(\xi_2)\end{vmatrix}.} \tag{8.51}$$

8.2.5.2 Wave Function of Three-Particle Systems

For a system of three noninteracting identical particles, the symmetric wave function is given by

$$\psi_s(\xi_1,\xi_2,\xi_3) = \frac{1}{\sqrt{3!}}\sum_P \hat{P}\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\psi_{n_3}(\xi_3), \tag{8.52}$$

or by

$$\begin{aligned}\psi_s(\xi_1,\xi_2,\xi_3) = \frac{1}{\sqrt{3!}}\Big[&\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\psi_{n_3}(\xi_3)+\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_3)\psi_{n_3}(\xi_2)\\ &+\psi_{n_1}(\xi_2)\psi_{n_2}(\xi_1)\psi_{n_3}(\xi_3)+\psi_{n_1}(\xi_2)\psi_{n_2}(\xi_3)\psi_{n_3}(\xi_1)\\ &+\psi_{n_1}(\xi_3)\psi_{n_2}(\xi_1)\psi_{n_3}(\xi_2)+\psi_{n_1}(\xi_3)\psi_{n_2}(\xi_2)\psi_{n_3}(\xi_1)\Big],\end{aligned} \tag{8.53}$$

and, when $n_1 \neq n_2 \neq n_3$, the antisymmetric wave function is given by

$$\psi_a(\xi_1,\xi_2,\xi_3) = \frac{1}{\sqrt{3!}}\sum_P (-1)^P \hat{P}\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\psi_{n_3}(\xi_3), \tag{8.54}$$

or, in the form of a determinant, by

$$\psi_a(\xi_1,\xi_2,\xi_3)=\frac{1}{\sqrt{3!}}\begin{vmatrix}\psi_{n_1}(\xi_1) & \psi_{n_1}(\xi_2) & \psi_{n_1}(\xi_3)\\ \psi_{n_2}(\xi_1) & \psi_{n_2}(\xi_2) & \psi_{n_2}(\xi_3)\\ \psi_{n_3}(\xi_1) & \psi_{n_3}(\xi_2) & \psi_{n_3}(\xi_3)\end{vmatrix}. \tag{8.55}$$

If $n_1=n_2=n_3=n$ we have $\psi_s(\xi_1,\xi_2,\xi_3)=\psi_n(\xi_1)\psi_n(\xi_2)\psi_n(\xi_3)$ and $\psi_a(\xi_1,\xi_2,\xi_3)=0$.

8.2.5.3 Wave Function of Many-Particle Systems

We can generalize (8.52) and (8.55) and write the symmetric and antisymmetric wave functions for a system of N noninteracting identical particles as follows:

$$\psi_s(\xi_1,\xi_2,\ldots,\xi_N) = \frac{1}{\sqrt{N!}}\sum_P \hat{P}\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\cdots\psi_{n_N}(\xi_N), \tag{8.56}$$

$$\psi_a(\xi_1,\xi_2,\ldots,\xi_N) = \frac{1}{\sqrt{N!}}\sum_P (-1)^P\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\cdots\psi_{n_N}(\xi_N), \tag{8.57}$$

or

$$\psi_a(\xi_1,\xi_2,\ldots,\xi_N)=\frac{1}{\sqrt{N!}}\begin{vmatrix}\psi_{n_1}(\xi_1) & \psi_{n_1}(\xi_2) & \cdots & \psi_{n_1}(\xi_N)\\ \psi_{n_2}(\xi_1) & \psi_{n_2}(\xi_2) & \cdots & \psi_{n_2}(\xi_N)\\ \vdots & \vdots & \ddots & \vdots\\ \psi_{n_N}(\xi_1) & \psi_{n_N}(\xi_2) & \cdots & \psi_{n_N}(\xi_N)\end{vmatrix}. \tag{8.58}$$

This $N\times N$ determinant, which involves one-particle states only, is known as the *Slater determinant*. An interchange of any pair of particles corresponds to an interchange of two columns of the determinant; this interchange introduces a change in the sign of the determinant. For even permutations we have $(-1)^P=1$, and for odd permutations we have $(-1)^P=-1$.

The relations (8.56) and (8.58) are valid for the case where the numbers $n_1, n_2, \ldots, n_N$ are all different from one another. What happens if some, or all, of these numbers are equal? In the symmetric case, if $n_1=n_2=\ldots=n_N$ then ψ_s is given by

$$\psi_s(\xi_1,\ \xi_2,\ldots,\ \xi_N)=\prod_{i=1}^{N}\psi_n(\xi_i)=\psi_n(\xi_1)\psi_n(\xi_2)\cdots\psi_n(\xi_N). \tag{8.59}$$

When there is a multiplicity in the numbers $n_1, n_2, \ldots, n_N$ (i.e., when some of the numbers n_i occur more than once), we have to be careful and avoid double counting. For instance, if n_1 occurs N_1 times in the sequence $n_1, n_2, \ldots, n_N$, if n_2 occurs N_2 times, and so on, the symmetric wave function will be given by

$$\psi_s(\xi_1,\ \xi_2,\ \ldots,\ \xi_N)=\sqrt{\frac{N_1!N_2!\cdots N_N!}{N!}}\sum_P \hat{P}\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\cdots\psi_{n_N}(\xi_N); \tag{8.60}$$

the summation $\sum_P$ is taken only over permutations which lead to *distinct* terms and includes $N!/N_1!N_2!\cdots N_n!$ different terms. For example, in the case of a system of three independent,

identical bosons where $n_1 = n_2 = n$ and $n_3 \neq n$, the multiplicity of n_1 is $N_1 = 2$; hence ψ_s is given by

$$\psi_s(\xi_1, \xi_2, \xi_3) = \sqrt{\frac{2!}{3!}} \sum_P \hat{P}\psi_n(\xi_1)\psi_n(\xi_2)\psi_{n_3}(\xi_3) = \frac{1}{\sqrt{3}} \Big[\psi_n(\xi_1)\psi_n(\xi_2)\psi_{n_3}(\xi_3) + \psi_n(\xi_1)\psi_{n_3}(\xi_2)\psi_n(\xi_3) + \psi_{n_3}(\xi_1)\psi_n(\xi_2)\psi_n(\xi_3) \Big]. \tag{8.61}$$

Unlike the symmetric case, the antisymmetric case is quite straightforward: if, among the numbers $n_1, n_2, \ldots, n_N$, only two are equal, the antisymmetric wave function vanishes. For instance, if $n_i = n_j$, the ith and jth rows of the determinant (8.58) will be identical; hence the determinant vanishes identically. Antisymmetric wave functions, therefore, are nonzero only for those cases where all the numbers $n_1, n_2, \ldots, n_N$ are different.

8.3 The Pauli Exclusion Principle

As mentioned above, if any two particles occupy the same single-particle state, the determinant (8.58), and hence the total wave function, will vanish since two rows of the determinant will be identical. We can thus infer that *in a system of N identical particles, no two fermions can occupy the same single-particle state at a time*; every single-particle state can be occupied by at most one fermion. This is the *Pauli exclusion principle*, which was first postulated in 1925 to explain the periodic table (see Chapter 1). It states that *no two electrons can occupy simultaneously the same (single-particle) quantum state on the same atom*; there can be only one (or at most one) electron occupying a state of four quantum numbers $n_i l_i m_{l_i} m_{s_i}$: $\psi_{n_i l_i m_{l_i} m_{s_i}}(\vec{r}_i, \vec{S}_i)$. The exclusion principle plays an important role in the structure of atoms. It has a direct effect on the *spatial distribution* of fermions.

Boson condensation

What about bosons? Do they have any restriction like fermions? Not at all. There is no restriction on the number of bosons that can occupy a single state. Instead of the exclusion principle of fermions, bosons tend to condense all in the same state, the ground state; this is called *boson condensation*. For instance, all the particles of liquid ^{4}He (a boson system) occupy the same ground state. This phenomenon is known as *Bose–Einstein condensation*. The properties of liquid ^{3}He are, however, completely different from those of liquid ^{4}He, because ^{3}He is a fermion system.

Remark

We have seen that when the Schrödinger equation involves the spin, the wave function of a single particle is equal to the product of the spatial part and the spin part: $\Psi(\vec{r}, \vec{S}) = \psi(\vec{r})\chi(\vec{S})$. The wave function of a system of N particles, which have spins, is the product of the spatial part and the spin part:

$$\Psi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2; \ldots; \vec{r}_N, \vec{S}_N) = \psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)\chi(\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_N). \tag{8.62}$$

This wave function must satisfy the appropriate symmetry requirements when the N particles are identical. In the case of a system of N identical bosons, the wave function must be

symmetric; hence the spatial and spin parts must have the same parity:

$$\Psi_s(\vec{r}_1,\vec{S}_1;\ \vec{r}_2,\vec{S}_2;\ \ldots;\ \vec{r}_N,\vec{S}_N) = \begin{cases} \psi_a(\vec{r}_1,\ \vec{r}_2,\ \ldots,\ \vec{r}_N)\chi_a(\vec{S}_1,\ \vec{S}_2,\ \ldots,\ \vec{S}_N), \\ \psi_s(\vec{r}_1,\ \vec{r}_2,\ \ldots,\ \vec{r}_N)\chi_s(\vec{S}_1,\ \vec{S}_2,\ \ldots,\ \vec{S}_N). \end{cases} \tag{8.63}$$

In the case of a system of N identical fermions, however, the space and spin parts must have different parities, leading to an overall wave function that is antisymmetric:

$$\Psi_a(\vec{r}_1,\vec{S}_1;\ \vec{r}_2,\vec{S}_2;\ \ldots;\ \vec{r}_N,\vec{S}_N) = \begin{cases} \psi_a(\vec{r}_1,\ \vec{r}_2,\ \ldots,\ \vec{r}_N)\chi_s(\vec{S}_1,\ \vec{S}_2,\ \ldots,\ \vec{S}_N), \\ \psi_s(\vec{r}_1,\ \vec{r}_2,\ \ldots,\ \vec{r}_N)\chi_a(\vec{S}_1,\ \vec{S}_2,\ \ldots,\ \vec{S}_N). \end{cases} \tag{8.64}$$

Example 8.3 (Wave function of two identical, noninteracting particles)
Find the wave functions of two systems of identical, noninteracting particles: the first consists of two bosons and the second of two spin $\frac{1}{2}$ fermions.

Solution
For a system of two identical, noninteracting bosons, (8.47) and (8.48) yield

$$\Psi_s(\vec{r}_1,\vec{S}_1;\vec{r}_2,\vec{S}_2) = \frac{1}{\sqrt{2}}\begin{cases} \left[\psi_{n_1}(\vec{r}_1)\psi_{n_2}(\vec{r}_2) - \psi_{n_1}(\vec{r}_2)\psi_{n_2}(\vec{r}_1)\right]\chi_a(\vec{S}_1,\ \vec{S}_2), \\ \left[\psi_{n_1}(\vec{r}_1)\psi_{n_2}(\vec{r}_2) + \psi_{n_1}(\vec{r}_2)\psi_{n_2}(\vec{r}_1)\right]\chi_s(\vec{S}_1,\ \vec{S}_2), \end{cases} \tag{8.65}$$

and for a system of two spin $\frac{1}{2}$ fermions

$$\Psi_a(\vec{r}_1,\vec{S}_1;\vec{r}_2,\vec{S}_2) = \frac{1}{\sqrt{2}}\begin{cases} \left[\psi_{n_1}(\vec{r}_1)\psi_{n_2}(\vec{r}_2) - \psi_{n_1}(\vec{r}_2)\psi_{n_2}(\vec{r}_1)\right]\chi_s(\vec{S}_1,\ \vec{S}_2), \\ \left[\psi_{n_1}(\vec{r}_1)\psi_{n_2}(\vec{r}_2) + \psi_{n_1}(\vec{r}_2)\psi_{n_2}(\vec{r}_1)\right]\chi_a(\vec{S}_1,\ \vec{S}_2), \end{cases} \tag{8.66}$$

where, from the formalism of angular momentum addition, there are three states (a triplet) that are symmetric, $\chi_s(\vec{S}_1,\ \vec{S}_2)$:

$$\chi_{triplet}(\vec{S}_1,\ \vec{S}_2) = \begin{cases} \left|\tfrac{1}{2}\ \tfrac{1}{2}\right\rangle_1 \left|\tfrac{1}{2}\ \tfrac{1}{2}\right\rangle_2, \\ \frac{1}{\sqrt{2}}\left(\left|\tfrac{1}{2}\ \tfrac{1}{2}\right\rangle_1 \left|\tfrac{1}{2}\ -\tfrac{1}{2}\right\rangle_2 + \left|\tfrac{1}{2}\ -\tfrac{1}{2}\right\rangle_1 \left|\tfrac{1}{2}\ \tfrac{1}{2}\right\rangle_2\right), \\ \left|\tfrac{1}{2}\ -\tfrac{1}{2}\right\rangle_1 \left|\tfrac{1}{2}\ -\tfrac{1}{2}\right\rangle_2, \end{cases} \tag{8.67}$$

and one state (a singlet) that is antisymmetric, $\chi_a(\vec{S}_1,\ \vec{S}_2)$:

$$\chi_{singlet}(\vec{S}_1,\ \vec{S}_2) = \frac{1}{\sqrt{2}}\left(\left|\frac{1}{2}\ \frac{1}{2}\right\rangle_1 \left|\frac{1}{2}\ -\frac{1}{2}\right\rangle_2 - \left|\frac{1}{2}\ -\frac{1}{2}\right\rangle_1 \left|\frac{1}{2}\ \frac{1}{2}\right\rangle_2\right). \tag{8.68}$$

8.4 The Exclusion Principle and the Periodic Table

Explaining the periodic table is one of the most striking successes of the Schrödinger equation. When combined with the Pauli exclusion principle, the equation offers insightful information on the structure of multielectron atoms.

In Chapter 6, we saw that the state of the hydrogen's electron, which moves in the spherically symmetric Coulomb potential of the nucleus, is described by four quantum numbers n, l, m_l, and m_s: $\Psi_{nlm_lm_s}(\vec{r}) = \psi_{nlm_l}(\vec{r})\chi_{m_s}$, where $\psi_{nlm_l}(\vec{r}) = R_{nl}(r)Y_{lm_l}(\theta,\varphi)$ is the electron's wave function when the spin is ignored and $\chi_{m_s} = \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle$ is the spin's state. This representation turns out to be suitable for any atom as well.

In a multielectron atom, the average potential in which every electron moves is different from the Coulomb potential of the nucleus; yet, to a good approximation, it can be assumed to be spherically symmetric. We can therefore, as in hydrogen, characterize the electronic states by the four quantum numbers n, l, m_l, and m_s, which respectively represent the principal quantum number, the orbital quantum number, the magnetic (or azimuthal) quantum number, and the spin quantum number; m_l represents the z-component of the electron orbital angular momentum and m_s the z-component of its spin.

Atoms have a *shell structure*. Each atom has a number of major shells that are specified by the radial or principal quantum number n. Shells have subshells which are specified by the orbital quantum number l. Subshells in turn have subsubshells, called *orbitals*, specified by m_l; so *an orbital is fully specified by three quantum numbers* n, l, m_l; i.e., it is defined by $|nlm_l\rangle$. Each shell n therefore has n subshells corresponding to $l = 0, 1, 2, 3, \ldots, n-1$, and in turn each subshell has $2l+1$ orbitals (or subsubshells), since to $m_l = -l, -l+1, -l+2, \ldots, l-2, l-1, l$. As in hydrogen, individual electrons occupy single-particle states or orbitals; the states corresponding to the respective numerical values $l = 0, 1, 2, 3, 4, 5, \ldots$ are called s, p, d, f, g, h, ... states. Hence for a given n an s-state has 1 orbital ($m_l = 0$), a p-state has 3 orbitals ($m_l = -1, 0, 1$), a d-state has 5 orbitals ($m_l = -2, -1, 0, 1, 2$), and so on (Chapter 6). We will label the electronic states by nl where, as before, l refers to s, p, d, f, etc.; for example 1s corresponds to $(n,l) = (1,0)$, 2s corresponds to $(n,l) = (2,0)$, 2p corresponds to $(n,l) = (2,1)$, 3s corresponds to $(n,l) = (3,0)$, and so on.

How do electrons fill the various shells and subshells in an atom? If electrons were bosons, they would all group in the ground state $|nlm_l\rangle = |100\rangle$; we wouldn't then have the rich diversity of elements that exist in nature. But since electrons are identical fermions, they are governed by the Pauli exclusion principle, which states that *no two electrons can occupy simultaneously the same quantum state* $|nlm_lm_s\rangle$ *on the same atom*. Hence each orbital state $|nlm_l\rangle$ can be occupied by two electrons *at most*: one having spin-up $m_s = +\frac{1}{2}$, the other spin-down $m_s = -\frac{1}{2}$. Hence, each state nl can accommodate $2(2l+1)$ electrons. So an s-state (i.e., $|n00\rangle$) can at most hold 2 electrons, a p-state (i.e., $|n1m_l\rangle$) at most 6 electrons, a d-state (i.e., $|n2m_l\rangle$) at most 10 electrons, an f-state (i.e., $|n3m_l\rangle$) at most 14 electrons, and so on (Figure 8.2).

For an atom in the ground state, the *electrons fill the orbitals in order of increasing energy*; once a subshell is filled, the next electron goes into the vacant subshell whose energy is just above the previous subshell. When all orbitals in a major electronic shell are filled up, we get a *closed shell*; the next electron goes into the next major shell, and so on. By filling the atomic orbitals one after the other in order of increasing energy, one obtains all the elements of the periodic table (Table 8.1).

Elements $1 \le Z \le 18$

As shown in Table 8.1, the first period (or first horizontal row) of the periodic table has two elements, hydrogen H and helium He; the second period has 8 elements, lithium Li to neon Ne; the third period also has 8 elements, sodium Na to argon Ar; and so on. The orbitals of the 18 lightest elements, $1 \le Z \le 18$, are filled in order of increasing energy according to the sequence: 1s, 2s, 2p, 3s, 3p. The electronic state of an atom is determined by specifying

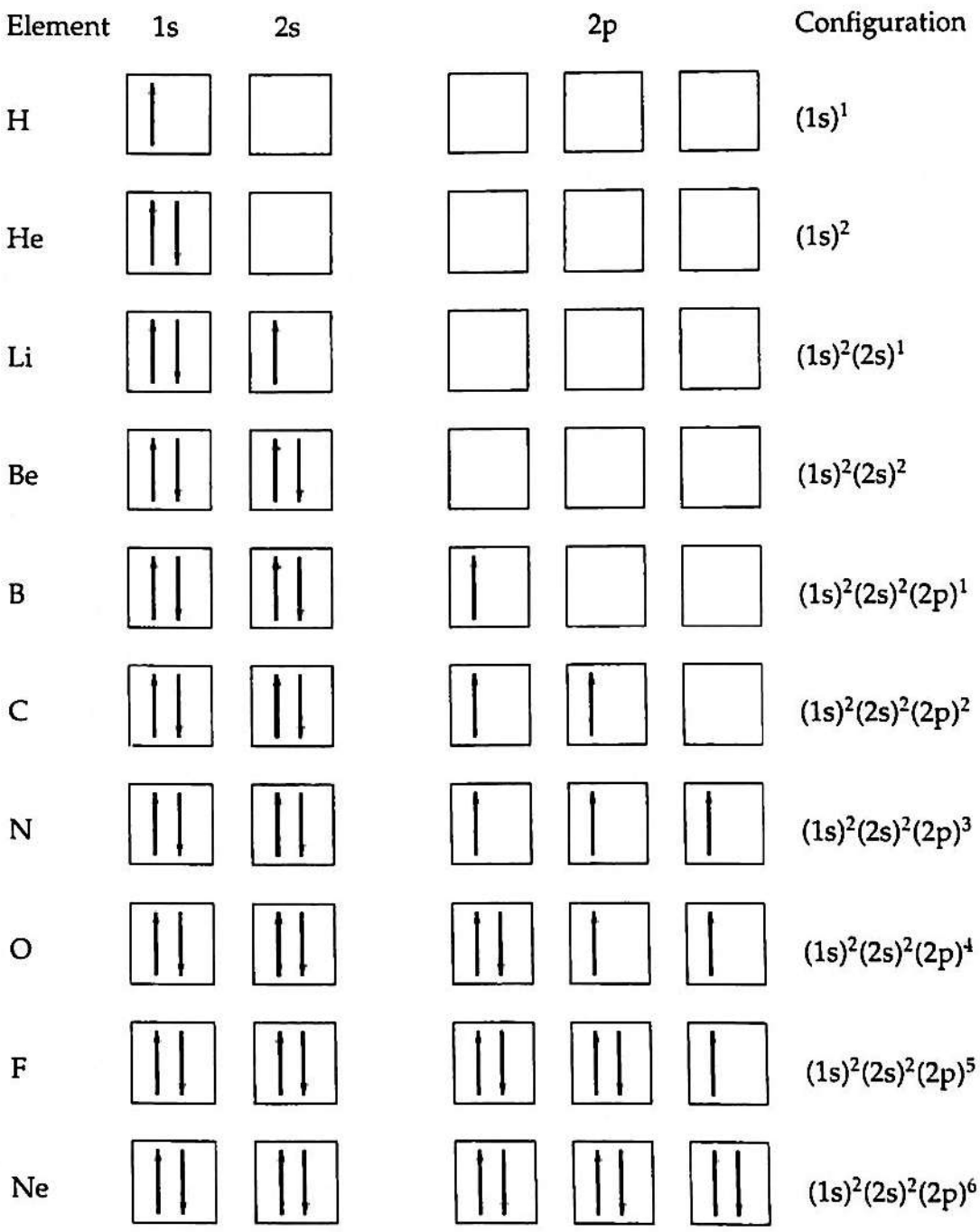


Figure 8.2: Filling orbitals according to the Pauli exclusion principle.

the occupied orbitals or by what is called the *electronic configuration*. For example, hydrogen has one electron, its ground state configuration is $(1s)^1$; helium He has two electrons: $(1s)^2$; lithium Li has three electrons: $(1s)^2(2s)^1$; beryllium Be has four: $(1s)^2(2s)^2$, and so on.

Now let us see how to determine the total angular momentum of an atom. For this, we need to calculate the total orbital angular momentum $\vec{L} = \sum_{i=1}^{Z} \vec{l}_i$, the total spin $\vec{S} = \sum_{i=1}^{Z} \vec{s}_i$, and then obtain total angular momentum by coupling $\vec{L}$ and $\vec{S}$, i.e., $\vec{J} = \vec{L} + \vec{S}$, where $\vec{l}_i$ and $\vec{s}_i$ are the orbital and spin angular momenta of individual electrons. As will be seen in Chapter 9, when the spin–orbit coupling is considered, the degeneracy of the atom's energy levels is partially lifted, introducing a splitting of the levels. The four numbers L, S, J, and M are good quantum numbers, where $|L - S| \leq J \leq L + S$ and $-J \leq M \leq J$. So there are $2S + 1$ values of J when $L \geq S$ and $2L + 1$ values when $L < S$. Since the energy depends on J, the levels corresponding to an L and S split into a $(2J + 1)$-multiplet. The issue now is to determine which one of these states has the lowest energy. Before studying this issue, let us introduce the spectroscopic notation according to which the state of an atom is labeled by

$$^{2S+1}L_J, \tag{8.69}$$

where, as before, the numbers $L = 0, 1, 2, 3, \ldots$ are designated by S, P, D, F, ..., respectively (we should mention here that the capital letters S, P, D, F, ... refer to the total orbital angular momentum of an atom, while the small letters s, p, d, f, ... refer to individual electrons; that is, s, p, d, f, ... describe the angular momentum states of individual electrons). For example, since the total angular momentum of a beryllium atom is $J = 0$, because $L = 0$ (all electrons are in s-states, $l_i = 0$) and $S = 0$ (both electrons in the $(1s)^2$ state are paired and so are the two electrons in the $(2s)^2$ state), the ground state of beryllium can be written as $^1S_0$. This applies actually to all other closed shell atoms such as helium He, neon Ne, argon Ar, and so on; their ground states are all specified by $^1S_0$ (Table 8.1).

Let us now consider boron B: the closed shells 1s and 2s have $L = S = J = 0$. Thus the angular momentum of boron is determined by the 1p electron which has $S = 1/2$ and $L = 1$. A coupling of $S = 1/2$ and $L = 1$ yields $J = 1/2$ or $3/2$, leading therefore to two possible states:

$$^2P_{1/2} \quad \text{or} \quad ^2P_{3/2}. \tag{8.70}$$

Which one has a lower energy? Before answering this question, let us consider another example, the carbon atom.

The ground state configuration of the carbon atom, as given by $(1s)^2(2s)^2(2p)^2$, implies that its total angular momentum is determined by the two 2p electrons. The coupling of the two spins $s = 1/2$, as shown in equations (7.182) to (7.185), yields two values for their total spin $S = 0$ or $S = 1$; and, as shown in Problem 7.3, page 481, a coupling of two individual orbital angular momenta $l = 1$ yields three values for the total angular momenta $L = 0$, 1, or 2. But the exclusion principle dictates that the total wave function has to be antisymmetric, i.e., the spin and orbital parts of the wave function must have opposite symmetries. Since the singlet spin state $S = 0$ is antisymmetric, the spin triplet $S = 1$ is symmetric, the orbital triplet $L = 1$ is antisymmetric, the orbital quintuplet $L = 2$ is symmetric, and the orbital singlet $L = 0$ is symmetric, the following states are antisymmetric:

$$^1S_0, \quad ^3P_0, \quad ^3P_1, \quad ^3P_2, \quad \text{or} \quad ^1D_2; \tag{8.71}$$

hence any one of these states can be the ground state of carbon. Again, which one of them has the lowest energy?

Table 8.1: Ground state electron configurations, spectroscopic description, and ionization energies for the first four rows of the periodic table. The brackets designate closed-shell elements.

Shell	Z	Element	Ground state configuration	Spectroscopic description	Ionization energy (eV)
1	1	H	$(1s)^1$	$^2S_{1/2}$	13.60
	2	He	$(1s)^2$	$^1S_0$	24.58
2	3	Li	$[He](2s)^1 = (1s)^2(2s)^1$	$^2S_{1/2}$	5.39
	4	Be	$[He](2s)^2$	$^1S_0$	9.32
	5	B	$[He](2s)^2(2p)^1$	$^2P_{1/2}$	8.30
	6	C	$[He](2s)^2(2p)^2$	$^3P_0$	11.26
	7	N	$[He](2s)^2(2p)^3$	$^4S_{3/2}$	14.55
	8	O	$[He](2s)^2(2p)^4$	$^3P_2$	13.61
	9	F	$[He](2s)^2(2p)^5$	$^2P_{3/2}$	17.42
	10	Ne	$[He](2s)^2(2p)^6$	$^1S_0$	21.56
3	11	Na	$[Ne](3s)^1$	$^2S_{1/2}$	5.14
	12	Mg	$[Ne](3s)^2$	$^1S_0$	7.64
	13	Al	$[Ne](3s)^2(3p)^1$	$^2P_{1/2}$	5.94
	14	Si	$[Ne](3s)^2(3p)^2$	$^3P_0$	8.15
	15	P	$[Ne](3s)^2(3p)^3$	$^4S_{3/2}$	10.48
	16	S	$[Ne](3s)^2(3p)^4$	$^3P_2$	10.36
	17	Cl	$[Ne](3s)^2(3p)^5$	$^2P_{3/2}$	13.01
	18	Ar	$[Ne](3s)^2(3p)^6$	$^1S_0$	15.76
4	19	K	$[Ar](4s)^1$	$^2S_{1/2}$	4.34
	20	Ca	$[Ar](4s)^2$	$^1S_0$	6.11
	21	Sc	$[Ar](3d)^1(4s)^2$	$^2D_{3/2}$	6.54
	22	Ti	$[Ar](3d)^2(4s)^2$	$^3F_2$	6.83
	23	V	$[Ar](3d)^3(4s)^2$	$^4F_{3/2}$	6.74
	24	Cr	$[Ar](3d)^5(4s)^1$	$^7S_3$	6.76
	25	Mn	$[Ar](3d)^5(4s)^2$	$^6S_{3/2}$	7.43
	26	Fe	$[Ar](3d)^6(4s)^2$	$^5D_4$	7.87
	27	Co	$[Ar](3d)^7(4s)^2$	$^4F_{9/2}$	7.86
	28	Ni	$[Ar](3d)^8(4s)^2$	$^3F_4$	7.63
	29	Cu	$[Ar](3d)^{10}(4s)^1$	$^2S_{1/2}$	7.72
	30	Zn	$[Ar](3d)^{10}(4s)^2$	$^1S_0$	9.39
	31	Ga	$[Ar](3d)^{10}(4s)^2(4p)^1$	$^2P_{1/2}$	6.00
	32	Ge	$[Ar](3d)^{10}(4s)^2(4p)^2$	$^3P_0$	7.88
	33	As	$[Ar](3d)^{10}(4s)^2(4p)^3$	$^4S_{3/2}$	9.81
	34	Se	$[Ar](3d)^{10}(4s)^2(4p)^4$	$^3P_2$	9.75
	35	Br	$[Ar](3d)^{10}(4s)^2(4p)^5$	$^2P_{3/2}$	11.84
	36	Kr	$[Ar](3d)^{10}(4s)^2(4p)^6$	$^1S_0$	9.81

To answer this question and the question pertaining to (8.70), we may invoke *Hund's rules*: (a) the lowest energy level corresponds to the state with the largest spin S (i.e., the maximum number of electrons have unpaired spins); (b) among the states with a given value of S, the lowest energy level corresponds to the state with the largest value of L; (c) for a subshell that is less than half full the lowest energy state corresponds to $J = |L - S|$, and for a subshell that is more than half full the lowest energy state corresponds to $J = L + S$.

Hund's third rule answers the question pertaining to (8.70): since the 2p shell of boron is less than half full, the value of J corresponding to the lowest energy is given by $J = |L - S| = 1 - 1/2 = 1/2$; hence $^2P_{1/2}$ is the lower energy state.

To find which one of the states (8.71) has the lowest energy, Hund's first rule dictates that $S = 1$. Since the triplet $S = 1$ is symmetric, we need an antisymmetric spatial wave function; this is given by the spatial triplet $L = 1$. We are thus left with three possible choices: $J = 0, 1$, or 2. Hund's third rule precludes the values $J = 1$ and 2. Since the 2p shell of carbon is less than half full, the value of J corresponding to the lowest energy is given by $J = |L - S| = 1 - 1 = 0$; hence $^3P_0$ is the lower energy state (Table 8.1). That is, the two electrons are in different spatial states or different orbitals (Figure 8.2). Actually, we could have guessed this result: since the Coulomb repulsion between the two electrons when they are paired together is much larger than when they are unpaired, the lower energy configuration corresponds to the case where the electrons are in different spatial states. The ground state configurations of the remaining elements, oxygen to argon, can be inferred in a similar way (Table 8.1).

Elements $Z \geq 18$

When the 3p shell is filled, one would expect to place the next electron in a 3d shell. But this doesn't take place due to the occurrence of an interesting effect: the 4s states have lower energy than the 3d states. Why? In a hydrogen atom the states 3s, 3p, and 3d have the same energy ($E_3^{(0)} = -\mathcal{R}/3^2 = -1.51\,\text{eV}$, since $\mathcal{R} = 13.6\,\text{eV}$). But in multielectron atoms, these states have different energy values. As l increases, the effective repulsive potential $\hbar^2 l(l+1)/2mr^2$ causes the d-state electrons to be thrown outward and the s-state electrons to remain closer to the nucleus. Being closer to the nucleus, the s-state electrons therefore feel the full attraction of the nucleus, whereas the d-state electrons experience a much weaker attraction. This is known as the *screening effect*, because the inner electrons, i.e., the s-state electrons, screen the nucleus; hence the outward electrons (the d-state electrons) do not experience the full attraction of the nucleus, but instead feel a weak effective potential. As a result, the energy of the 3d-state is larger than that of the 4s-state. The screening effect also causes the energy of the 5s-state to have a lower energy than the 4d-state, and so on. So for a given n, the energies E_{nl} increase as l increases; in fact, neglecting the spin–orbit interaction and considering relativistic corrections we will show in Chapter 9 (9.91) that the ground state energy depends on the principal and orbital quantum numbers n and l as $E_{nl}^{(0)} = Z^2 E_n^{(0)}\{1+\alpha^2 Z^2[2/(2l+1)-3/4n]/n\}$, where $\alpha = 1/137$ is the fine structure constant and $E_n^{(0)} = -\mathcal{R}/n^2 = -13.6\,\text{eV}/n^2$.

In conclusion, the periodic table can be obtained by filling the orbitals in order of increasing energy E_{nl} as follows (Table 8.1):

$$1s^2,\ 2s^2,\ 2p^6,\ 3s^2,\ 3p^6,\ 4s^2,\ 3d^{10},\ 4p^6,\ 5s^2,\ 4d^{10},\ 5p^6,\ 6s^2,\ 4f^{14},$$
$$5d^{10},\ 6p^6,7s^2,\ 5f^{14},\ 6d^{10},\ 7p^6,\ldots. \tag{8.72}$$

Remarks

The chemical properties of an element is mostly determined by the outermost shell. Hence elements with similar electron configurations for the outside shell have similar chemical properties. This is the idea behind the structure of the periodic table: it is arranged in a way that all elements in a column have similar chemical properties. For example, the elements in the last column, helium, neon, argon, krypton, and so on, have the outer p-shell completely filled (except for helium whose outside shell is 1s). These atoms, which are formed when a shell or a subshell is filled, are very stable, interact very weakly with one another, and do not combine with other elements to form molecules or new compounds; that is, they are chemically inert. They are very reluctant to give up or to accept an electron. Due to these properties, they are called *noble gases*. They have a very low boiling point (around −200 °C). Note that each row of the periodic table corresponds to filling out a shell or subshell of the atom, up to the next noble gas. Also, there is a significant energy gap before the next level is encountered after each of these elements. As shown in Table 8.1, a large energy is required to ionize these elements; for instance, 24.58 eV is needed to ionize a helium atom.

Atoms consisting of a closed shell (or a noble gas configuration) plus an s-electron (or a valence electron), such as Li, Na, K, and so on, have the lowest binding energy; these elements are known as the alkali metals. In elements consisting of an alkali configuration plus an electron, the second s-electron is more bound than the valence electron of the alkali atom because of the higher nuclear charge. As the p-shell is gradually filled (beyond the noble gas configuration), the binding energy increases initially (as in boron, carbon, and nitrogen) till the fourth electron, then it begins to drop (Table 8.1). This is due to the fact that when the p-shell is less than half full all spins are parallel; hence all three spatial wave functions are antisymmetric. With the fourth electron (as in oxygen), two spins will be antiparallel or paired; hence the spatial wave function is not totally antisymmetric, causing a drop in the energy. Note that elements with one electron more than or one electron less than noble gas configurations are the most active chemically, because they tend to easily give up or easily accept one electron.

Example 8.4

(a) Specify the total angular momenta corresponding to ^{4}G, ^{3}H, and ^{1}D.

(b) Find the spectroscopic notation for the ground state configurations of aluminum Al ($Z = 13$) and scandium Sc ($Z = 21$).

Solution

(a) For the term ^{4}G the orbital angular momentum is $L = 4$ and the spin is $S = 3/2$, since $2S + 1 = 4$. The values of the total angular momentum corresponding to the coupling of $L = 4$ and $S = 3/2$ are given by $|4 - 3/2| \leq J \leq 4 + 3/2$. Hence we have $J = 5/2, 7/2, 9/2, 11/2$.

Similarly, for ^{3}H we have $S = 1$ and $L = 5$. Therefore, we have $|5 - 1| \leq J \leq 5 + 1$, or $J = 4, 5, 6$.

For ^{1}D we have $S = 0$ and $L = 2$. Therefore, we have $|2 - 0| \leq J \leq 2 + 0$, or $J = 2$.

(b) The ground state configuration of Al is $[\text{Ne}](3s)^2(3p)^1$. The total angular momentum of this element is determined by the 3p electron, because $S = 0$ and $L = 0$ for both [Ne] and $(3s)^2$. Since the 3p electron has $S = 1/2$ and $L = 1$, the total angular momentum is given by $|1 - 1/2| \leq J \leq 1 + 1/2$. Hence we have $J = 1/2, 3/2$. Which of the values $J = 1/2$ and $J = 3/2$ has a lower energy? According to Hund's third rule, since the 3p shell is less than half full, the state $J = |L - S| = 1/2$ has the lower energy. Hence the ground state configuration of Al corresponds to $^2P_{1/2}$ (Table 8.1), where we have used the spectroscopic notation $^{2S+1}L_J$.

Since the ground state configuration of Sc is $[Ar](4s)^2(3d)^1$, the angular momentum is given by that of the 3d electron. Since $S = 1/2$ and $L = 2$, and since the 3d shell is less than half full, Hund's third rule dictates that the total angular momentum is given by $J = |L - S| = |2 - 1/2| = 3/2$. Hence we have $^2D_{3/2}$.

8.5 Solved Problems

Problem 8.1
Consider a system of three noninteracting particles that are confined to move in an infinite square-well potential well of length a: $V(x) = 0$ for $0 \le x \le a$ and $V(x) = +\infty$ for other values of x. Determine the energy and wave function of the ground state and the first and second excited states when the three particles are (a) spinless and distinguishable with $m_1 < m_2 < m_3$; (b) identical bosons; (c) identical spin $\frac{1}{2}$ particles; and (d) distinguishable spin $\frac{1}{2}$ particles.

Solution
(a) As shown in Example 8.2 on page 507, the total energy and wave function are given by

$$E_{n_1,n_2,n_3} = \frac{\hbar^2\pi^2}{2a^2}\left(\frac{n_1^2}{m_1} + \frac{n_2^2}{m_2} + \frac{n_3^2}{m_3}\right), \tag{8.73}$$

$$\psi_{n_1,n_2,n_3}(x_1, x_2, x_3) = \sqrt{\frac{8}{a^3}}\sin\left(\frac{n_1\pi}{a}x_1\right)\sin\left(\frac{n_2\pi}{a}x_2\right)\sin\left(\frac{n_3\pi}{a}x_3\right). \tag{8.74}$$

The ground state of the system corresponds to the case where all three particles occupy their respective ground state orbitals, $n_1 = n_2 = n_3 = 1$; hence

$$E^{(0)} = E_{1,1,1} = \frac{\hbar^2\pi^2}{2a^2}\left(\frac{1}{m_1} + \frac{1}{m_2} + \frac{1}{m_3}\right), \tag{8.75}$$

$$\psi^{(0)}(x_1, x_2, x_3) = \psi_{1,1,1}(x_1, x_2, x_3) = \sqrt{\frac{8}{a^3}}\sin\left(\frac{\pi}{a}x_1\right)\sin\left(\frac{\pi}{a}x_2\right)\sin\left(\frac{\pi}{a}x_3\right). \tag{8.76}$$

Since particle 3 has the highest mass, the first excited state corresponds to the case where particle 3 is in $n_3 = 2$, while particles 1 and 2 remain in $n_1 = n_2 = 1$:

$$E^{(1)} = E_{1,1,2} = \frac{\hbar^2\pi^2}{2a^2}\left(\frac{1}{m_1} + \frac{1}{m_2} + \frac{4}{m_3}\right), \tag{8.77}$$

$$\psi^{(1)}(x_1, x_2, x_3) = \psi_{1,1,2}(x_1, x_2, x_3) = \sqrt{\frac{8}{a^3}}\sin\left(\frac{\pi}{a}x_1\right)\sin\left(\frac{\pi}{a}x_2\right)\sin\left(\frac{2\pi}{a}x_3\right). \tag{8.78}$$

Similarly, the second excited state corresponds to the case where particles 1 and 3 are in $n_1 = n_3 = 1$, while particle 2 is in $n_2 = 2$:

$$E^{(2)} = E_{1,2,1} = \frac{\hbar^2\pi^2}{2a^2}\left(\frac{1}{m_1} + \frac{4}{m_2} + \frac{1}{m_3}\right), \tag{8.79}$$

$$\psi^{(2)}(x_1, x_2, x_3) = \psi_{1,2,1}(x_1, x_2, x_3) = \sqrt{\frac{8}{a^3}} \sin\left(\frac{\pi}{a}x_1\right)\sin\left(\frac{2\pi}{a}x_2\right)\sin\left(\frac{\pi}{a}x_3\right). \tag{8.80}$$

(b) If all three particles are identical bosons, the ground state will correspond to all particles in the lowest state $n_1 = n_2 = n_3 = 1$ (Figure 8.3):

$$E^{(0)} = E_{1,1,1} = 3\varepsilon_1 = \frac{3\hbar^2\pi^2}{2ma^2}, \tag{8.81}$$

$$\psi^{(0)} = \psi_1(x_1)\psi_1(x_2)\psi_1(x_3) = \sqrt{\frac{8}{a^3}} \sin\left(\frac{\pi}{a}x_1\right)\sin\left(\frac{\pi}{a}x_2\right)\sin\left(\frac{\pi}{a}x_3\right), \tag{8.82}$$

since $\psi_n(x_i) = \sqrt{2/a}\sin(n\pi x_i/a)$.

In the first excited state we have two particles in ψ_1 (each with energy $\varepsilon_1 = \hbar^2\pi^2/(2ma^2)$) and one in ψ_2 (with energy $\varepsilon_2 = 4\hbar^2\pi^2/(2ma^2) = 4\varepsilon_1$):

$$E^{(1)} = 2\varepsilon_1 + \varepsilon_2 = 2\varepsilon_1 + 4\varepsilon_1 = 6\varepsilon_1 = \frac{3\pi^2\hbar^2}{ma^2}. \tag{8.83}$$

The wave function is somewhat tricky again. Since the particles are identical, we can no longer say which particle is in which state; all we can say is that two particles are in ψ_1 and one in ψ_2. Since the value $n = 1$ occurs twice (two particles are in ψ_1), we infer from (8.60) and (8.61) that

$$\psi^{(1)}(x_1, x_2, x_3) = \sqrt{\frac{2!}{3!}}\Big[\psi_1(x_1)\psi_1(x_2)\psi_2(x_3) + \psi_1(x_1)\psi_2(x_2)\psi_1(x_3) + \psi_2(x_1)\psi_1(x_2)\psi_1(x_3)\Big]. \tag{8.84}$$

In the second excited state we have one particle in ψ_1 and two in ψ_2:

$$E^{(2)} = \varepsilon_1 + 2\varepsilon_2 = \varepsilon_1 + 8\varepsilon_1 = 9\varepsilon_1 = \frac{9\pi^2\hbar^2}{2ma^2}. \tag{8.85}$$

Now, since the value $n = 2$ occurs twice (two particles are in ψ_2) and $n = 1$ only once, (8.60) and (8.61) yield

$$\psi^{(2)}(x_1, x_2, x_3) = \sqrt{\frac{2!}{3!}}\Big[\psi_1(x_1)\psi_2(x_2)\psi_2(x_3) + \psi_2(x_1)\psi_1(x_2)\psi_2(x_3) + \psi_2(x_1)\psi_2(x_2)\psi_1(x_3)\Big]. \tag{8.86}$$

(c) If the three particles were identical spin $\frac{1}{2}$ fermions, the ground state corresponds to the case where two particles are in the lowest state ψ_1 (one having a spin-up $|+\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle$, the other with a spin-down $|-\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle$), while the third particle is in the next state ψ_2 (its spin can be either up or down, $|\pm\rangle = \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle$); see Figure 8.3. The ground state energy is

$$E^{(0)} = 2\varepsilon_1 + \varepsilon_2 = 2\varepsilon_1 + 4\varepsilon_1 = 6\varepsilon_1 = \frac{3\hbar^2\pi^2}{ma^2}. \tag{8.87}$$

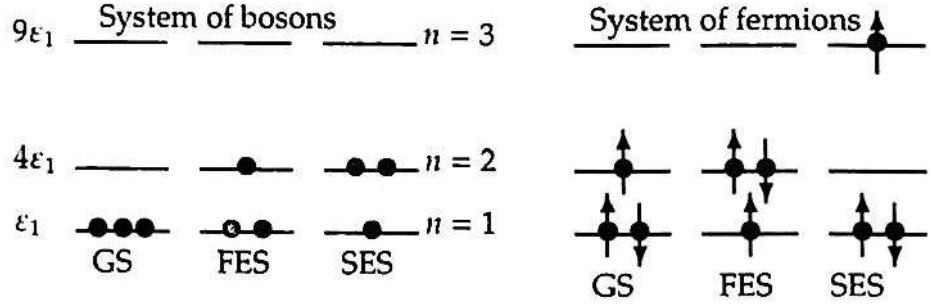


Figure 8.3: Particle distribution among the levels of the ground state (GS) and the first (FES) and second excited states (SES) for a system of three noninteracting identical bosons (left) and fermions (right) moving in an infinite well, with $\varepsilon_1 = \hbar^2\pi^2/(2ma^2)$. Each state of the fermion system is fourfold degenerate due to the various possible orientations of the spins.

The ground state wave function is antisymmetric and, in accordance with (8.55), it is given by

$$\psi^{(0)}(x_1, x_2, x_3) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \psi_1(x_1)\chi(S_1) & \psi_1(x_2)\chi(S_2) & \psi_1(x_3)\chi(S_3) \\ \psi_1(x_1)\chi(S_1) & \psi_1(x_2)\chi(S_2) & \psi_1(x_3)\chi(S_3) \\ \psi_2(x_1)\chi(S_1) & \psi_2(x_2)\chi(S_2) & \psi_2(x_3)\chi(S_3) \end{vmatrix}. \tag{8.88}$$

This state is fourfold degenerate, since there are four different ways of configuring the spins of the three fermions (the ground state (GS) shown in Figure 8.3 is just one of the four configurations). **Remark:** one should be careful not to erroneously conclude that, since the first and second rows of the determinant in (8.88) are "identical", the determinant is zero. We should keep in mind that the spin states are given by $\chi(S_1) = |\pm\rangle$, $\chi(S_2) = |\pm\rangle$, and $\chi(S_3) = |\pm\rangle$; hence, we need to select these spin states in such a way that no two rows (nor two columns) of the determinant are identical. For instance, one of the possible configurations of the ground state wave function is given by

$$\psi^{(0)}(x_1, x_2, x_3) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \psi_1(x_1)\,|+\rangle & \psi_1(x_2)\,|-\rangle & \psi_1(x_3)\,|+\rangle \\ \psi_1(x_1)\,|-\rangle & \psi_1(x_2)\,|+\rangle & \psi_1(x_3)\,|+\rangle \\ \psi_2(x_1)\,|+\rangle & \psi_2(x_2)\,|+\rangle & \psi_2(x_3)\,|-\rangle \end{vmatrix}. \tag{8.89}$$

This remark applies also to the first and second excited state wave functions (8.90) and (8.92); it also applies to the wave function (8.109).

The first excited state corresponds to one particle in the lowest state ψ_1 (its spin can be either up or down) and the other two particles in the state ψ_2 (the spin of one is up, the other is down). As in the ground state, there are also four different ways of configuring the spins of the three fermions in the first excited state (FES); the FES shown in Figure 8.3 is just one of the four configurations:

$$\psi^{(1)}(x_1, x_2, x_3) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \psi_1(x_1)\chi(S_1) & \psi_1(x_2)\chi(S_2) & \psi_1(x_3)\chi(S_3) \\ \psi_2(x_1)\chi(S_1) & \psi_2(x_2)\chi(S_2) & \psi_2(x_3)\chi(S_3) \\ \psi_2(x_1)\chi(S_1) & \psi_2(x_2)\chi(S_2) & \psi_2(x_3)\chi(S_3) \end{vmatrix}. \tag{8.90}$$

These four different states correspond to the same energy

$$E^{(1)} = \varepsilon_1 + 2\varepsilon_2 = \varepsilon_1 + 8\varepsilon_1 = 9\varepsilon_1 = \frac{9\hbar^2\pi^2}{2ma^2}. \tag{8.91}$$

The excitation energy of the first excited state is $E^{(1)} - E^{(0)} = 9\varepsilon_1 - 6\varepsilon_1 = 3\hbar^2\pi^2/(2ma^2)$.

The second excited state corresponds to two particles in the lowest state ψ_1 (one with spin-up, the other with spin-down) and the third particle in the third state ψ_3 (its spin can be either up or down). This state also has four different spin configurations; hence it is fourfold degenerate:

$$\psi^{(2)}(x_1, x_2, x_3) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \psi_1(x_1)\chi(S_1) & \psi_1(x_2)\chi(S_2) & \psi_1(x_3)\chi(S_3) \\ \psi_1(x_1)\chi(S_1) & \psi_1(x_2)\chi(S_2) & \psi_1(x_3)\chi(S_3) \\ \psi_3(x_1)\chi(S_1) & \psi_3(x_2)\chi(S_2) & \psi_3(x_3)\chi(S_3) \end{vmatrix}. \tag{8.92}$$

The energy of the second excited state is

$$E^{(2)} = 2\varepsilon_1 + \varepsilon_3 = 2\varepsilon_1 + 9\varepsilon_1 = 11\varepsilon_1 = \frac{11\hbar^2\pi^2}{2ma^2}. \tag{8.93}$$

The excitation energy of this state is $E^{(2)} - E^{(0)} = 11\varepsilon_1 - 6\varepsilon_1 = 5\varepsilon_1 = 5\hbar^2\pi^2/(2ma^2)$.

(d) If the particles were distinguishable fermions, there will be no restrictions on the symmetry of the wave function, neither on the space part nor on the spin part. The values of the energy of the ground state, the first excited state, and the second excited state will be similar to those calculated in part (a). However, the wave functions of these states are somewhat different from those found in part (a); while the states derived in (a) are nondegenerate, every state of the current system is eightfold degenerate, since the coupling of three $\frac{1}{2}$ spins yield eight different spin states (Chapter 7). So the wave functions of the system are obtained by multiplying each of the space wave functions $\psi^{(0)}(x_1, x_2, x_3)$, $\psi^{(1)}(x_1, x_2, x_3)$, and $\psi^{(2)}(x_1, x_2, x_3)$, derived in (a), by any of the eight spin states calculated in Chapter 7:

$$\left|1, \frac{3}{2}, \pm\frac{3}{2}\right\rangle = \left|\frac{1}{2}, \frac{1}{2}, \frac{1}{2}; \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}\right\rangle, \tag{8.94}$$

$$\left|1, \frac{3}{2}, \pm\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\Big(\mid j_1, j_2, j_3; \mp, \pm, \pm\rangle + \mid j_1, j_2, j_3; \pm, \mp, \pm\rangle + \mid j_1, j_2, j_3; \pm, \pm, \mp\rangle\Big), \tag{8.95}$$

$$\left|0, \frac{1}{2}, \pm\frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}}\Big(\mid j_1, j_2, j_3; \pm, \pm, \mp\rangle - \mid j_1, j_2, j_3; \mp, \pm, \pm >\Big), \tag{8.96}$$

$$\left|1, \frac{1}{2}, \pm\frac{1}{2}\right\rangle = \frac{\mp 1}{\sqrt{6}}\Big(\mid j_1, j_2, j_3; \pm, \pm, \mp\rangle - 2 \mid j_1, j_2, j_3; \pm, \mp, \pm\rangle + \mid j_1, j_2, j_3; \mp, \pm, \pm\rangle\Big). \tag{8.97}$$

Problem 8.2

Consider a system of three noninteracting identical spin $\frac{1}{2}$ particles that are in the same spin state $\left|\frac{1}{2}, \frac{1}{2}\right\rangle$ and confined to move in an infinite square-well potential well of length a: $V(x) = 0$ for $0 \le x \le a$ and $V(x) = +\infty$ for other values of x. Determine the energy and wave function of the ground state, the first excited state, and the second excited state.

Solution

We may mention first that the single-particle energy and wave function of a particle moving in an infinite well are given by $\varepsilon_n = n^2\hbar^2\pi^2/(2ma^2)$ and $\psi_n(x_i) = \sqrt{2/a}\sin(n\pi x_i/a)$.

The wave function of this system is antisymmetric, since it consists of identical fermions. Moreover, since all the three particles are in the same spin state, no two particles can be in the same state; every energy level is occupied by at most one particle. For instance, the ground state corresponds to the case where the three lowest levels $n = 1,2,3$ are occupied by one particle each. The ground state energy and wave function are thus given by

$$E^{(0)} = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = \varepsilon_1 + 4\varepsilon_1 + 9\varepsilon_1 = 14\varepsilon_1 = \frac{7\hbar^2\pi^2}{ma^2}, \tag{8.98}$$

$$\psi^{(0)}(x_1,x_2,x_3) = \frac{1}{\sqrt{3!}}\begin{vmatrix} \psi_1(x_1) & \psi_1(x_2) & \psi_1(x_3) \\ \psi_2(x_1) & \psi_2(x_2) & \psi_2(x_3) \\ \psi_3(x_1) & \psi_3(x_2) & \psi_3(x_3) \end{vmatrix} \left|\frac{1}{2}, \frac{1}{2}\right\rangle. \tag{8.99}$$

The first excited state is obtained (from the ground state) by raising the third particle to the fourth level: the levels $n = 1$, 2, and 4 are occupied by one particle each and the third level is empty:

$$E^{(1)} = \varepsilon_1 + \varepsilon_2 + \varepsilon_4 = \varepsilon_1 + 4\varepsilon_1 + 16\varepsilon_1 = 21\varepsilon_1 = \frac{21\hbar^2\pi^2}{2ma^2}, \tag{8.100}$$

$$\psi^{(1)}(x_1,x_2,x_3) = \frac{1}{\sqrt{3!}}\begin{vmatrix} \psi_1(x_1) & \psi_1(x_2) & \psi_1(x_3) \\ \psi_2(x_1) & \psi_2(x_2) & \psi_2(x_3) \\ \psi_4(x_1) & \psi_4(x_2) & \psi_4(x_3) \end{vmatrix} \left|\frac{1}{2}, \frac{1}{2}\right\rangle. \tag{8.101}$$

In the second excited state, the levels $n = 1,3,4$ are occupied by one particle each; the second level is empty:

$$E^{(2)} = \varepsilon_1 + \varepsilon_3 + \varepsilon_4 = \varepsilon_1 + 9\varepsilon_1 + 16\varepsilon_1 = 26\varepsilon_1 = \frac{13\hbar^2\pi^2}{ma^2}, \tag{8.102}$$

$$\psi^{(2)}(x_1,x_2,x_3) = \frac{1}{\sqrt{3!}}\begin{vmatrix} \psi_1(x_1) & \psi_1(x_2) & \psi_1(x_3) \\ \psi_3(x_1) & \psi_3(x_2) & \psi_3(x_3) \\ \psi_4(x_1) & \psi_4(x_2) & \psi_4(x_3) \end{vmatrix} \left|\frac{1}{2}, \frac{1}{2}\right\rangle. \tag{8.103}$$

Problem 8.3

Consider a system of N noninteracting identical particles that are confined to move in an infinite square-well potential well of length a: $V(x) = 0$ for $0 \le x \le a$ and $V(x) = +\infty$ for other values of x. Find the ground state energy and wave function of this system when the particles are (a) bosons and (b) spin $\frac{1}{2}$ fermions.

Solution

In the case of a particle moving in an infinite well, its energy and wave function are $\varepsilon_n = n^2\hbar^2\pi^2/(2ma^2)$ and $\psi_n(x_i) = \sqrt{2/a}\sin(n\pi x_i/a)$.

(a) In the case where the N particles are bosons, the ground state is obtained by putting all the particles in the state $n = 1$; the energy and wave function are then given by

$$E^{(0)} = \varepsilon_1 + \varepsilon_1 + \varepsilon_1 + \cdots + \varepsilon_1 = N\varepsilon_1 = \frac{N\hbar^2\pi^2}{2ma^2}, \tag{8.104}$$

$$\psi^{(0)}(x_1,x_2,\ldots,x_N) = \prod_{i=1}^{N}\sqrt{\frac{2}{a}}\sin\left(\frac{\pi}{a}x_i\right) = \sqrt{\frac{2^N}{a^N}}\sin\left(\frac{\pi}{a}x_1\right)\sin\left(\frac{\pi}{a}x_2\right)\cdots\sin\left(\frac{\pi}{a}x_N\right). \tag{8.105}$$

(b) In the case where the N particles are spin $\frac{1}{2}$ fermions, each level can be occupied by at most two particles having different spin states $\left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle$. The ground state is thus obtained by distributing the N particles among the $N/2$ lowest levels at a rate of two particles per level:

$$E^{(0)} = 2\varepsilon_1 + 2\varepsilon_2 + 2\varepsilon_3 + \cdots + 2\varepsilon_{N/2} = 2\sum_{n=1}^{N/2} \frac{n^2\hbar^2\pi^2}{2ma^2} = \frac{\hbar^2\pi^2}{ma^2}\sum_{n=1}^{N/2} n^2. \tag{8.106}$$

If N is large we may calculate $\sum_{n=1}^{N/2} n^2$ by using the approximation

$$\sum_{n=1}^{N/2} n^2 \simeq \int_1^{N/2} n^2 dn \simeq \frac{1}{3}\left(\frac{N}{2}\right)^3; \tag{8.107}$$

hence the ground state energy will be given by

$$E^{(0)} \simeq N^3 \frac{\hbar^2\pi^2}{24ma^2}. \tag{8.108}$$

The average energy per particle is $E^{(0)}/N \simeq N^2\hbar^2\pi^2/(24ma^2)$. In the case where N is even, a possible configuration of the ground state wave function $\psi^{(0)}(x_1, x_2, \ldots, x_N)$ is given as follows:

$$\frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(x_1)\chi(S_1) & \psi_1(x_2)\chi(S_2) & \cdots & \psi_1(x_N)\chi(S_N) \\ \psi_1(x_1)\chi(S_1) & \psi_1(x_2)\chi(S_2) & \cdots & \psi_1(x_N)\chi(S_N) \\ \psi_2(x_1)\chi(S_1) & \psi_2(x_2)\chi(S_2) & \cdots & \psi_2(x_N)\chi(S_N) \\ \psi_2(x_1)\chi(S_1) & \psi_2(x_2)\chi(S_2) & \cdots & \psi_2(x_N)\chi(S_N) \\ \psi_3(x_1)\chi(S_1) & \psi_3(x_2)\chi(S_2) & \cdots & \psi_3(x_N)\chi(S_N) \\ \psi_3(x_1)\chi(S_1) & \psi_3(x_2)\chi(S_2) & \cdots & \psi_3(x_N)\chi(S_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{N/2}(x_1)\chi(S_1) & \psi_{N/2}(x_2)\chi(S_2) & \cdots & \psi_{N/2}(x_N)\chi(S_N) \\ \psi_{N/2}(x_1)\chi(S_1) & \psi_{N/2}(x_2)\chi(S_2) & \cdots & \psi_{N/2}(x_N)\chi(S_N) \end{vmatrix}, \tag{8.109}$$

where $\chi(S_i) = \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle$ is the spin state of the ith particle, with $i = 1, 2, 3, \ldots, N$. If N is odd then we need to remove the last row of the determinant.

Problem 8.4

Neglecting the spin–orbit interaction and the interaction between the electrons, find the energy levels and the wave functions of the three lowest states for a two-electron atom.

Solution

Examples of such a system are the helium atom (Z = 2), the singly ionized Li^+ ion (Z = 3), the doubly ionized Be^{2+} ion (Z = 4), and so on. Neglecting the spin–orbit interaction and the interaction between the electrons, $V_{12} = e^2/r_{12} = e^2/|\vec{r}_1 - \vec{r}_2|$, we can view each electron as moving in the Coulomb field of the Ze nucleus. The Hamiltonian of this system is therefore equal to the sum of the Hamiltonians of the two electrons:

$$\hat{H} = H_0^{(1)} + H_0^{(2)} = \left(-\frac{\hbar^2}{2\mu}\nabla_1^2 - \frac{Ze^2}{r_1}\right) + \left(-\frac{\hbar^2}{2\mu}\nabla_2^2 - \frac{Ze^2}{r_2}\right), \tag{8.110}$$

where $\mu = Mm_e/(M+m_e)$, M is the mass of the nucleus, and m_e is the mass of the electron. We have considered here that the nucleus is placed at the origin and that the electrons are located at $\vec{r}_1$ and $\vec{r}_2$. The Schrödinger equation of the system is given by

$$\left[\hat{H}_0^{(1)} + \hat{H}_0^{(2)}\right]\Psi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = E_{n_1 n_2}\Psi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2), \tag{8.111}$$

where the energy $E_{n_1 n_2}$ is equal to the sum of the energies of the electrons:

$$E_{n_1 n_2} = E_{n_1}^{(0)} + E_{n_2}^{(0)} = -\frac{Z^2 e^2}{2a_0}\frac{1}{n_1^2} - \frac{Z^2 e^2}{2a_0}\frac{1}{n_2^2}, \tag{8.112}$$

where $a_0 = \hbar^2/(me^2)$ is the Bohr radius. The wave function is equal to the product of the spatial and spin parts:

$$\Psi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = \psi(\vec{r}_1, \vec{r}_2)\chi(\vec{S}_1, \vec{S}_2); \tag{8.113}$$

$\vec{S}_1$ and $\vec{S}_2$ are the spin vectors of the electrons.

Since this system consists of two identical fermions (electrons), its wave function has to be antisymmetric. So either the spatial part is antisymmetric and the spin part is symmetric,

$$\Psi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = \frac{1}{\sqrt{2}}\left[\phi_{n_1 l_1 m_1}(\vec{r}_1)\phi_{n_2 l_2 m_2}(\vec{r}_2) - \phi_{n_2 l_2 m_2}(\vec{r}_1)\phi_{n_1 l_1 m_1}(\vec{r}_2)\right]\chi_{triplet}(\vec{S}_1, \vec{S}_2), \tag{8.114}$$

or the spatial part is symmetric and the spin part is antisymmetric,

$$\Psi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = \phi_{n_1 l_1 m_1}(\vec{r}_1)\phi_{n_2 l_2 m_2}(\vec{r}_2)\chi_{singlet}(\vec{S}_1, \vec{S}_2), \tag{8.115}$$

where $\chi_{triplet}$ and $\chi_{singlet}$, which result from the coupling of two spins $\frac{1}{2}$, are given by (8.67) and (8.68).

Let us now specify the energy levels and wave functions of the three lowest states. The ground state corresponds to both electrons occupying the lowest state $|nlm> = |100>$ (i.e., $n_1 = n_2 = 1$); its energy and wave function can be inferred from (8.112) and (8.115):

$$E^{(0)} = E_{11} = 2E_1^{(0)} = -2\frac{Z^2 e^2}{2a_0} = -27.2Z^2\,\text{eV}, \tag{8.116}$$

$$\Psi_0(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = \phi_{100}(\vec{r}_1)\phi_{100}(\vec{r}_2)\chi_{singlet}(\vec{S}_1, \vec{S}_2), \tag{8.117}$$

where $\phi_{100}(\vec{r}) = R_{10}(r)Y_{00}(\Omega) = (1/\sqrt{\pi})(Z/a_0)^{3/2}e^{-Zr/a_0}$.

In the first excited state, one electron occupies the lowest level $|nlm\rangle = |100\rangle$ and the other electron occupies the level $|nlm\rangle = |200\rangle$; this corresponds either to $n_1 = 1$, $n_2 = 2$ or to $n_1 = 2$, $n_2 = 1$. The energy and the wave function can thus be inferred from (8.112) and (8.114):

$$E^{(1)} = E_{12} = E_1^{(0)} + E_2^{(0)} = -\frac{Z^2 e^2}{2a_0} - \frac{1}{4}\frac{Z^2 e^2}{2a_0} = -\frac{5}{4} \times 13.6Z^2\,\text{eV} = -17.0Z^2\,\text{eV}, \tag{8.118}$$

$$\Psi_1(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = \frac{1}{\sqrt{2}}\left[\phi_{100}(\vec{r}_1)\phi_{200}(\vec{r}_2) - \phi_{200}(\vec{r}_1)\phi_{100}(\vec{r}_2)\right]\chi_{triplet}(\vec{S}_1, \vec{S}_2), \tag{8.119}$$

where $\phi_{200}(\vec{r}) = R_{20}(r)Y_{00}(\Omega) = (1/\sqrt{8\pi})(Z/a_0)^{3/2}(1 - Zr/2a_0)e^{-Zr/2a_0}$.

Finally, the energy and wave function of the second excited state, which correspond to both electrons occupying the second level $|nlm\rangle = |200\rangle$ (i.e., $n_1 = n_2 = 2$), can be inferred from (8.112) and (8.115):

$$E^{(2)} = E_{22} = E_2^{(0)} + E_2^{(0)} = 2E_2^{(0)} = -\frac{1}{2}\frac{Z^2e^2}{2a_0} = -\frac{1}{2} \times 13.6Z^2\,\text{eV} = -6.8Z^2\,\text{eV}, \tag{8.120}$$

$$\Psi_2(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = \phi_{200}(\vec{r}_1)\phi_{200}(\vec{r}_2)\chi_{singlet}(\vec{S}_1, \vec{S}_2). \tag{8.121}$$

These results are obviously not expected to be accurate because, by neglecting the Coulomb interaction between the electrons, we have made a grossly inaccurate approximation. For instance, the numerical value for the ground state energy obtained from (8.112) for the helium atom is $E^{(0)}_{theory} = -108.8\,\text{eV}$ whereas the experimental value is $E^{(0)}_{exp} = -78.975\,\text{eV}$; that is, the theoretical value is 37.8% lower than the experimental value.

In Chapter 9 we will show how to use perturbation theory and the variational method to obtain very accurate theoretical values for the energy levels of two-electron atoms.

Problem 8.5

Find the energy levels and the wave functions of the ground state and the first excited state for a system of two noninteracting identical particles moving in a common external harmonic oscillator potential for (a) two spin 1 particles with no orbital angular momentum and (b) two spin $\frac{1}{2}$ particles.

Solution

Since the particles are noninteracting and identical, their Hamiltonian is $\hat{H} = \hat{H}_1 + \hat{H}_2$, where $\hat{H}_1$ and $\hat{H}_2$ are the Hamiltonians of particles 1 and 2: $\hat{H}_j = -(\hbar^2/2m)d^2/dx_j^2 + m\omega x_j^2/2$ with $j = 1, 2$. The total energy of the system is $E_{n_1n_2} = \varepsilon_{n_1} + \varepsilon_{n_2}$, where $\varepsilon_{n_j} = \left(n_j + \frac{1}{2}\right)\hbar\omega$.

(a) When the system consists of two identical spin 1 particles, the total wave function of this system must be symmetric. Thus, the space and spin parts must be both symmetric or both antisymmetric:

$$\Psi(x_1, S_1; x_2, S_2) = \frac{1}{\sqrt{2}}\Big[\psi_s(x_1, x_2)\chi_s(S_1, S_2) + \psi_a(x_1, x_2)\chi_a(S_1, S_2)\Big], \tag{8.122}$$

where

$$\psi_s(x_1, x_2) = \frac{1}{\sqrt{2}}\Big[\psi_{n_1}(x_1)\psi_{n_2}(x_2) + \psi_{n_1}(x_2)\psi_{n_2}(x_1)\Big], \tag{8.123}$$

$$\psi_a(x_1, x_2) = \frac{1}{\sqrt{2}}\left[\psi_{n_1}(x_1)\psi_{n_2}(x_2) - \psi_{n_1}(x_2)\psi_{n_2}(x_1)\right], \tag{8.124}$$

where $\psi_n(x)$ is a harmonic oscillator wave function for the state n; for instance, the ground state and first excited state are

$$\psi_0(x) = \frac{1}{\sqrt{\sqrt{\pi}x_0}}\exp\left(-\frac{x^2}{2x_0^2}\right), \qquad \psi_1(x) = \sqrt{\frac{2}{\sqrt{\pi}x_0^3}}x\exp\left(-\frac{x^2}{2x_0^2}\right), \tag{8.125}$$

with $x_0 = \sqrt{\hbar/(m\omega)}$.

The spin states $\chi(S_1, S_2)$ can be obtained by coupling the spins of the two particles, $S_1 = 1$ and $S_2 = 1$: $\vec{S} = \vec{S}_1 + \vec{S}_2$. As shown in Chapter 7, the spin states corresponding to $S = 2$ are given by

$$|2,\ \pm 2\rangle = |11;\ \pm 1, \pm 1\rangle, \qquad |2,\ \pm 1\rangle = \frac{1}{\sqrt{2}}\Big(\,|1,1;\ \pm 1, 0\rangle + |1,1;\ 0, \pm 1,\rangle\Big), \tag{8.126}$$

$$|2,\ 0\rangle = \frac{1}{\sqrt{6}}\Big(\,|1,1;\ 1,\ -1\rangle + 2|1,1;\ 0,0\rangle + |1,1;\ -1,1\rangle\Big), \tag{8.127}$$

those corresponding to $S = 1$ by

$$|1,\ \pm 1\rangle\ =\ \frac{1}{\sqrt{2}}\Big(\ \pm|1,1;\ \pm 1, 0\rangle \mp |1,1;\ 0, \pm 1\rangle\Big), \tag{8.128}$$

$$|1,\ 0\rangle\ =\ \frac{1}{\sqrt{2}}\Big(\,|1,1;\ 1,-1\rangle - |1,1;\ -1,1\rangle\Big), \tag{8.129}$$

and the one corresponding to $S = 0$ by

$$|0,\ 0\rangle = \frac{1}{\sqrt{3}}\Big(\,|1,1;\ 1,-1\rangle - |1,1;\ 0,0\rangle + |1,1;\ -1,1\rangle\Big). \tag{8.130}$$

Obviously, the five states $|2,\ m_s\rangle$, corresponding to $S = 2$ and $|00\rangle$, are symmetric, whereas the three states $|1,\ m_s\rangle$ are antisymmetric. Thus, $\chi_s(S_1, S_2)$ is given by any one of the six states $|2,\ \pm 2\rangle$, $|2,\ \pm 1\rangle$, $|2,\ 0\rangle$, and $|0,\ 0\rangle$; as for $\chi_a(S_1, S_2)$, it is given by any one of the three states $|2,\ \pm 1\rangle$, and $|1,\ 0\rangle$.

The ground state corresponds to the case where both particles are in their respective ground states $n_1 = n_2 = 0$. The energy is then given by $E^{(0)} = \varepsilon_0 + \varepsilon_0 = \frac{1}{2}\hbar\omega + \frac{1}{2}\hbar\omega = \hbar\omega$. Since $\psi_a(x_1, x_2)$, as given by (8.124), vanishes for $n_1 = n_2 = 0$, the ground state wave function (8.122) reduces to

$$\Psi_0(x_1, S_1; x_2, S_2) = \psi_0(x_1)\psi_0(x_2)\chi_s(S_1, S_2) = \frac{1}{\sqrt{\pi}x_0}\exp\left(-\frac{x_1^2 + x_2^2}{2x_0^2}\right)\chi_s(S_1, S_2), \tag{8.131}$$

where $\psi_0(x)$ is given by (8.125). The ground state is thus sixfold degenerate, since there are six spin states $\chi_s(S_1, S_2)$ that are symmetric.

In the first excited state, one particle occupies the ground state level $n = 0$ and the other in the first excited state $n = 1$; this corresponds to two possible configurations: either $n_1 = 0$ and $n_2 = 1$ or $n_1 = 1$ and $n_2 = 0$. The energy is then given by $E^{(1)} = \varepsilon_0 + \varepsilon_1 = \frac{1}{2}\hbar\omega + \frac{3}{2}\hbar\omega = 2\hbar\omega$. The first excited state can be inferred from (8.122) to (8.124):

$$\begin{aligned}\Psi_1(x_1, S_1; x_2, S_2)\ =\ & \frac{1}{2}\Big[\,\psi_0(x_1)\psi_1(x_2) + \psi_0(x_2)\psi_1(x_1)\Big]\chi_s(S_1, S_2) \\ & + \frac{1}{2}\Big[\,\psi_0(x_1)\psi_1(x_2) - \psi_0(x_2)\psi_1(x_1)\Big]\chi_a(S_1, S_2),\end{aligned} \tag{8.132}$$

where $\psi_0(x)$ and $\psi_1(x)$ are listed in (8.125). The first excited state is ninefold degenerate since there are six spin states, $\chi_s(S_1, S_2)$, that are symmetric and three, $\chi_a(S_1, S_2)$, that are antisymmetric.

(b) For a system of two identical fermions, the wave function must be antisymmetric and the space and spin parts must have opposite symmetries:

$$\Psi(x_1,S_1;x_2,S_2)=\frac{1}{2}\Big[\psi_s(x_1,x_2)\chi_{singlet}(S_1,S_2)+\psi_a(x_1,x_2)\chi_{triplet}(S_1,S_2)\Big], \tag{8.133}$$

where the symmetric spin state, $\chi_{triplet}(S_1,S_2)$, is given by the triplet states listed in (8.67); the antisymmetric spin state, $\chi_{singlet}(S_1,S_2)$, is given by the (singlet) state (8.68).

The ground state for the two spin $\frac{1}{2}$ particles corresponds to the case where both particles occupy the lowest level, $n_1=n_2=0$, and have different spin states. The energy is then given by $E^{(0)}=\varepsilon_0+\varepsilon_0=\hbar\omega$ and the wave function by

$$\begin{aligned}\Psi_0(x_1,S_1;x_2,S_2) &= \psi_0(x_1)\psi_0(x_2)\chi_{singlet}(S_1,S_2)\\ &= \frac{1}{\sqrt{\pi}x_0}\exp\left(-\frac{x_1^2+x_2^2}{2x_0^2}\right)\chi_{singlet}(S_1,S_2),\end{aligned} \tag{8.134}$$

since $\psi_a(x_1,x_2)$ vanishes for $n_1=n_2=0$. The ground state is not degenerate, since there is only one spin state which is antisymmetric, $\chi_{singlet}(S_1,S_2)$.

The first excited state corresponds also to $n_1=0$ and $n_2=1$ or $n_1=1$ and $n_2=0$. The energy is then given by $E^{(1)}=\varepsilon_0+\varepsilon_1=2\hbar\omega$ and the wave function by

$$\begin{aligned}\Psi_1(x_1,S_1;x_2,S_2) &= \frac{1}{2}\Big[\psi_0(x_1)\psi_1(x_2)+\psi_0(x_2)\psi_1(x_1)\Big]\chi_{singlet}(S_1,S_2)\\ &\quad+\frac{1}{2}\Big[\psi_0(x_1)\psi_1(x_2)-\psi_0(x_2)\psi_1(x_1)\Big]\chi_{triplet}(S_1,S_2).\end{aligned} \tag{8.135}$$

This state is fourfold degenerate since there are three spin states, $\chi_{triplet}(S_1,S_2)$, that are symmetric and one, $\chi_{singlet}(S_1,S_2)$, that is antisymmetric.

8.6 Exercises

Exercise 8.1
Consider a system of three noninteracting identical bosons that move in a common external one-dimensional harmonic oscillator potential. Find the energy levels and wave functions of the ground state, the first excited state, and the second excited state of the system.

Exercise 8.2
Consider two identical particles of spin $\frac{1}{2}$ that are confined in a cubical box of side L. Find the energy and the wave function of this system in the case of no interaction between the particles.

Exercise 8.3
(a) Consider a system of two nonidentical particles, each of spin 1 and having no orbital angular momentum (i.e., both particles are in s states). Write down all possible states for this system.

(b) What restrictions do we get if the two particles are identical? Write down all possible states for this system of two spin 1 identical particles.

Exercise 8.4
Two identical particles of spin $\frac{1}{2}$ are enclosed in a one-dimensional box potential of length L with rigid walls at $x = 0$ and $x = L$. Assuming that the two-particle system is in a *triplet spin state*, find the energy levels, the wave functions, and the degeneracies corresponding to the three lowest states.

Exercise 8.5
Two identical particles of spin $\frac{1}{2}$ are enclosed in a one-dimensional box potential of length L with rigid walls at $x = 0$ and $x = L$. Assuming that the two-particle system is in a *singlet spin state*, find the energy levels, the wave functions, and the degeneracies corresponding to the three lowest states.

Exercise 8.6
Two identical particles of spin $\frac{1}{2}$ are moving under the influence of a one-dimensional harmonic oscillator potential. Assuming that the two-particle system is in a *triplet spin state*, find the energy levels, the wave functions, and the degeneracies corresponding to the three lowest states.

Exercise 8.7
Find the ground state energy, the average ground state energy per particle, and the ground state wave function of a system of N noninteracting, identical bosons moving under the influence of a one-dimensional harmonic oscillator potential.

Exercise 8.8
Find the ground state energy, the average ground state energy per particle, and the ground state wave function of a system of N noninteracting identical spin $\frac{1}{2}$ particles moving under the influence of a one-dimensional harmonic oscillator potential for the following two cases:
(a) when N is even and
(b) when N is odd.

Exercise 8.9
Consider a system of four noninteracting particles that are confined to move in a one-dimensional infinite potential well of length a: $V(x) = 0$ for $0 < x < a$ and $V(x) = \infty$ for other values of x. Determine the energies and wave functions of the ground state, the first excited state, and the second excited state when the four particles are
(a) distinguishable bosons such that their respective masses satisfy this relation: $m_1 < m_2 < m_3 < m_4$, and
(b) identical bosons (each of mass m).

Exercise 8.10
Consider a system of four noninteracting identical spin 1/2 particles (each of mass m) that are confined to move in a one-dimensional infinite potential well of length a: $V(x) = 0$ for $0 < x < a$ and $V(x) = \infty$ for other values of x. Determine the energies and wave functions of the ground state and the first three excited states. Draw a figure showing how the particles are distributed among the levels.

Exercise 8.11
Consider a system of four noninteracting identical spin $\frac{1}{2}$ particles that are in the same spin

state $\left|\frac{1}{2}, \frac{1}{2}\right\rangle$ and confined to move in a one-dimensional infinite potential well of length a: $V(x) = 0$ for $0 < x < a$ and $V(x) = \infty$ for other values of x. Determine the energies and wave functions of the ground state, the first excited state, and the second excited state.

Exercise 8.12
Assuming the electrons in the helium atom to be spinless bosons and neglecting the interactions between them, find the energy and the wave function of the ground state and the first excited state of this (hypothetical) system.

Exercise 8.13
Assuming the electrons in the lithium atom to be spinless bosons and neglecting the interactions between them, find the energy and the wave function of the ground state and the first excited state of this (hypothetical) system.

Exercise 8.14
Consider a system of two noninteracting identical spin 1/2 particles (with mass m) that are confined to move in a one-dimensional infinite potential well of length L: $V(x) = 0$ for $0 < x < L$ and $V(x) = \infty$ for other values of x. Assume that the particles are in a state with the wave function

$$\Psi(x_1, x_2) = \frac{\sqrt{2}}{L}\left[\sin\left(\frac{2\pi x_1}{L}\right)\sin\left(\frac{5\pi x_2}{L}\right) + \sin\left(\frac{5\pi x_1}{L}\right)\sin\left(2\frac{\pi x_2}{L}\right)\right]\chi(s_1, s_2),$$

where x_1 and x_2 are the positions of particles 1 and 2, respectively, and $\chi(s_1, s_2)$ is the spin state of the two particles.

(a) Is $\chi(s_1, s_2)$ going to be a singlet or triplet state?
(b) Find the energy of this system.

Exercise 8.15
Consider a system of two noninteracting identical spin 1/2 particles (with mass m) that are confined to move in a common one-dimensional harmonic oscillator potential. Assume that the particles are in a state with the wave function

$$\Psi(x_1, x_2) = \frac{\sqrt{2}}{\sqrt{\pi}x_0^2}(x_2 - x_1)\exp\left(-\frac{x_1^2 + x_2^2}{2x_0^2}\right)\chi(s_1, s_2),$$

where x_1 and x_2 are the positions of particles 1 and 2, respectively, and $\chi(s_1, s_2)$ is the spin state of the two particles.

(a) Is $\chi(s_1, s_2)$ going to be a singlet or triplet state?
(b) Find the energy of this system.

Exercise 8.16
Consider a system of five noninteracting electrons (in the approximation where the Coulomb interaction between the electrons is neglected) that are confined to move in a common one-dimensional infinite potential well of length $L = 0.5$ nm: $V(x) = 0$ for $0 < x < L$ and $V(x) = \infty$ for other values of x.

(a) Find the ground state energy of the system.
(b) Find the energy of the first state of the system.
(c) Find the excitation energy of the first excited state.

Exercise 8.17
Determine the ground state electron configurations for the atoms having $Z = 40$, 53, 70, and 82 electrons.

Exercise 8.18
Specify the possible J values (i.e., total angular momenta) associated with each of the following states: ${}^1P$, ${}^4F$, ${}^2G$, and ${}^1H$.

Exercise 8.19
Find the spectroscopic notation ${}^{2S+1}L_J$ (i.e., find the L, S, and J) for the ground state configurations of

(a) Sc ($Z = 21$) and
(b) Cu ($Z = 29$).

Chapter 9

Approximation Methods for Stationary States

9.1 Introduction

Most problems encountered in quantum mechanics cannot be solved exactly. *Exact* solutions of the Schrödinger equation exist only for a few *idealized* systems. To solve general problems, one must resort to approximation methods. A variety of such methods have been developed, and each has its own area of applicability. In this chapter we consider approximation methods that deal with *stationary* states corresponding to *time-independent* Hamiltonians. In the following chapter we will deal with approximation methods for explicitly time-dependent Hamiltonians.

To study problems of stationary states, we focus on three approximation methods: *perturbation theory, the variational method,* and *the WKB method.*

Perturbation theory is based on the assumption that the problem we wish to solve is, in some sense, only *slightly* different from a problem that can be solved exactly. In the case where the deviation between the two problems is *small,* perturbation theory is suitable for calculating the contribution associated with this deviation; this contribution is then added as a correction to the energy and the wave function of the exactly solvable Hamiltonian. So perturbation theory builds on the known exact solutions to obtain approximate solutions.

What about those systems whose Hamiltonians cannot be reduced to an exactly solvable part plus a small correction? For these, we may consider the variational method or the WKB approximation. The variational method is particularly useful in estimating the energy eigenvalues of the ground state and the first few excited states of a system for which one has only a qualitative idea about the form of the wave function.

The WKB method is useful for finding the energy eigenvalues and wave functions of systems for which the *classical limit is valid.* Unlike perturbation theory, the variational and WKB methods do not require the existence of a closely related Hamiltonian that can be solved exactly.

The application of the approximation methods to the study of stationary states consists of finding the energy eigenvalues E_n and the eigenfunctions $|\psi_n\rangle$ of a time-independent

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

Hamiltonian $\hat{H}$ that does not have exact solutions:

$$\hat{H} \mid \psi_n\rangle = E_n \mid \psi_n\rangle. \tag{9.1}$$

Depending on the structure of $\hat{H}$, we can use any of the three methods mentioned above to find the approximate solutions to this eigenvalue problem.

9.2 Time-Independent Perturbation Theory

This method is most suitable when $\hat{H}$ is very close to a Hamiltonian $\hat{H}_0$ that can be solved exactly. In this case, $\hat{H}$ can be split into two time-independent parts

$$\hat{H} = \hat{H}_0 + \hat{H}_p, \tag{9.2}$$

where $\hat{H}_p$ is very small compared to $\hat{H}_0$ ($\hat{H}_0$ is known as the Hamiltonian of the unperturbed system). As a result, $\hat{H}_p$ is called the perturbation, for its effects on the energy spectrum and eigenfunctions will be small; such perturbation is encountered, for instance, in systems subject to *weak* electric or magnetic fields. We can make this idea more explicit by writing $\hat{H}_p$ in terms of a dimensionless real parameter λ which is very small compared to 1:

$$\hat{H}_p = \lambda \hat{W} \qquad (\lambda \ll 1). \tag{9.3}$$

Thus the eigenvalue problem (9.1) becomes

$$(\hat{H}_0 + \lambda \hat{W}) \mid \psi_n\rangle = E_n \mid \psi_n\rangle. \tag{9.4}$$

In what follows we are going to consider two separate cases depending on whether the exact solutions of $\hat{H}_0$ are nondegenerate or degenerate. Each of these two cases requires its own approximation scheme.

9.2.1 Nondegenerate Perturbation Theory

In this section we limit our study to the case where $\hat{H}_0$ has no degenerate eigenvalues; that is, for every energy $E_n^{(0)}$ there corresponds only one eigenstate $\mid \phi_n\rangle$:

$$\hat{H}_0 \mid \phi_n\rangle = E_n^{(0)} \mid \phi_n\rangle, \tag{9.5}$$

where the exact eigenvalues $E_n^{(0)}$ and exact eigenfunctions $\mid \phi_n\rangle$ are known.

The main idea of perturbation theory consists in assuming that the perturbed eigenvalues and eigenstates can both be expanded in power series in the parameter λ:

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots, \tag{9.6}$$

$$\mid \psi_n\rangle = \mid \phi_n\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \cdots. \tag{9.7}$$

We need to make two remarks. First, one might think that whenever the perturbation is sufficiently weak, the expansions (9.6) and (9.7) always exist. Unfortunately, this is not always the case. There are cases where the perturbation is small, yet E_n and $\mid \psi_n\rangle$ are not expandable in

powers of λ. Second, the series (9.6) and (9.7) are frequently not convergent. However, when λ is small, the first few terms do provide a reliable description of the system. So in practice, we keep only one or two terms in these expansions; hence the problem of nonconvergence of these series is avoided (we will deal later with the problem of convergence). Note that when $\lambda = 0$ the expressions (9.6) and (9.7) yield the unperturbed solutions: $E_n = E_n^{(0)}$ and $|\psi_n\rangle = |\phi_n\rangle$. The parameters $E_n^{(k)}$ and the kets $|\psi_n^{(k)}\rangle$ represent the kth corrections to the eigenenergies and eigenvectors, respectively.

The job of perturbation theory reduces then to the calculation of $E_n^{(1)}$, $E_n^{(2)}$, ... and $|\psi_n^{(1)}\rangle$, $|\psi_n^{(2)}\rangle$, ... In this section we shall be concerned only with the determination of $E_n^{(1)}$, $E_n^{(2)}$, and $|\psi_n^{(1)}\rangle$. Assuming that the unperturbed states $|\phi_n\rangle$ are nondegenerate, and substituting (9.6) and (9.7) into (9.4), we obtain

$$\begin{aligned}\left(\hat{H}_0 + \lambda\hat{W}\right)\left(|\phi_n\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \cdots\right) \\ = \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots\right)\left(|\phi_n\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \cdots\right).\end{aligned} \tag{9.8}$$

The coefficients of successive powers of λ on both sides of this equation must be equal. Equating the coefficients of the first three powers of λ, we obtain these results:

- Zero order in λ:

$$\hat{H}_0|\phi_n\rangle = E_n^{(0)}|\phi_n\rangle, \tag{9.9}$$

- First order in λ:

$$\hat{H}_0|\psi_n^{(1)}\rangle + \hat{W}|\phi_n\rangle = E_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(1)}|\phi_n\rangle, \tag{9.10}$$

- Second order in λ:

$$\hat{H}_0|\psi_n^{(2)}\rangle + \hat{W}|\psi_n^{(1)}\rangle = E_n^{(0)}|\psi_n^{(2)}\rangle + E_n^{(1)}|\psi_n^{(1)}\rangle + E_n^{(2)}|\phi_n\rangle. \tag{9.11}$$

We now proceed to determine the eigenvalues $E_n^{(1)}$, $E_n^{(2)}$ and the eigenvector $|\psi_n^{(1)}\rangle$ from (9.9) to (9.11). For this, we need to specify how the states $|\phi_n\rangle$ and $|\psi_n\rangle$ overlap. Since $|\psi_n\rangle$ is considered not to be very different from $|\phi_n\rangle$, we have $\langle\phi_n|\psi_n\rangle \simeq 1$. We can, however, normalize $|\psi_n\rangle$ so that its overlap with $|\phi_n\rangle$ is exactly equal to one:

$$\langle\phi_n|\psi_n\rangle = 1. \tag{9.12}$$

Substituting (9.7) into (9.12) we get

$$\lambda\langle\phi_n|\psi_n^{(1)}\rangle + \lambda^2\langle\phi_n|\psi_n^{(2)}\rangle + \cdots = 0; \tag{9.13}$$

hence the coefficients of the various powers of λ must vanish separately:

$$\langle\phi_n|\psi_n^{(1)}\rangle = \langle\phi_n|\psi_n^{(2)}\rangle = \cdots = 0. \tag{9.14}$$

First-order correction

To determine the first-order correction, $E_n^{(1)}$, to E_n we need simply to multiply both sides of (9.10) by $\langle \phi_n \mid$:

$$\boxed{E_n^{(1)} = \langle \phi_n \mid \hat{W} \mid \phi_n \rangle,} \tag{9.15}$$

where we have used the facts that $\langle \phi_n \mid \hat{H}_0 \mid \psi_n^{(1)} \rangle$ and $\langle \phi_n \mid \psi_n^{(1)} \rangle$ are both equal to zero and $\langle \phi_n \mid \phi_n \rangle = 1$. The insertion of (9.15) into (9.6) thus yields the energy to first-order perturbation:

$$\boxed{E_n = E_n^{(0)} + \langle \phi_n \mid \hat{H}_p \mid \phi_n \rangle.} \tag{9.16}$$

Note that for some systems, the first-order correction $E_n^{(1)}$ vanishes exactly. In such cases, one needs to consider higher-order terms.

Let us now determine $\mid \psi_n^{(1)} \rangle$. Since the set of the unperturbed states $\mid \phi_n \rangle$ form a complete and orthonormal basis, we can expand $\mid \psi_n^{(1)} \rangle$ in the $\{\mid \phi_n \rangle\}$ basis:

$$\mid \psi_n^{(1)} \rangle = \left(\sum_m \mid \phi_m \rangle \langle \phi_m \mid \right) \mid \psi_n^{(1)} \rangle = \sum_{m \neq n} \langle \phi_m \mid \psi_n^{(1)} \rangle \mid \phi_m \rangle; \tag{9.17}$$

the term $m = n$ does not contribute, since $\langle \phi_n \mid \psi_n^{(1)} \rangle = 0$. The coefficient $\langle \phi_m \mid \psi_n^{(1)} \rangle$ can be inferred from (9.10) by multiplying both sides by $\langle \phi_m \mid$:

$$\langle \phi_m \mid \psi_n^{(1)} \rangle = \frac{\langle \phi_m \mid \hat{W} \mid \phi_n \rangle}{E_n^{(0)} - E_m^{(0)}}, \tag{9.18}$$

which, when substituted into (9.17), leads to

$$\mid \psi_n^{(1)} \rangle = \sum_{m \neq n} \frac{\langle \phi_m \mid \hat{W} \mid \phi_n \rangle}{E_n^{(0)} - E_m^{(0)}} \mid \phi_m \rangle. \tag{9.19}$$

The eigenfunction $\mid \psi_n \rangle$ of $\hat{H}$ to first order in $\lambda \hat{W}$ can then be obtained by substituting (9.19) into (9.7):

$$\boxed{\mid \psi_n \rangle = \mid \phi_n \rangle + \sum_{m \neq n} \frac{\langle \phi_m \mid \hat{H}_p \mid \phi_n \rangle}{E_n^{(0)} - E_m^{(0)}} \mid \phi_m \rangle.} \tag{9.20}$$

Second-order correction

Now, to determine $E_n^{(2)}$ we need to multiply both sides of (9.11) by $\langle \phi_n \mid$:

$$E_n^{(2)} = \langle \phi_n \mid \hat{W} \mid \psi_n^{(1)} \rangle; \tag{9.21}$$

in obtaining this result we have used the facts that $\langle \phi_n \mid \psi_n^{(1)} \rangle = \langle \phi_n \mid \psi_n^{(2)} \rangle = 0$ and $\langle \phi_n \mid \phi_n \rangle = 1$. Inserting (9.19) into (9.21) we end up with

$$\boxed{E_n^{(2)} = \sum_{m \neq n} \frac{\left| \langle \phi_m \mid \hat{W} \mid \phi_n \rangle \right|^2}{E_n^{(0)} - E_m^{(0)}}.} \tag{9.22}$$

The eigenenergy to second order in $\hat{H}_p$ is obtained by substituting (9.22) and (9.15) into (9.6):

$$\boxed{E_n = E_n^{(0)} + \langle \phi_n \mid \hat{H}_p \mid \phi_n \rangle + \sum_{m \neq n} \frac{\left| \langle \phi_m | \hat{H}_p \mid \phi_n \rangle \right|^2}{E_n^{(0)} - E_m^{(0)}} + \cdots .} \tag{9.23}$$

In principle one can obtain energy corrections to any order. However, pushing the calculations beyond the second order, besides being mostly intractable, is a futile exercise, since the first two orders are generally sufficiently accurate.

Validity of the time-independent perturbation theory

For perturbation theory to work, the corrections it produces must be small; convergence must be achieved with the first two corrections. Expressions (9.20) and (9.23) show that the expansion parameter is $\langle \phi_m | \hat{H}_p \mid \phi_n \rangle / (E_n^{(0)} - E_m^{(0)})$. Thus, for the perturbation schemes (9.6) and (9.7) to work (i.e., to converge), the expansion parameter must be small:

$$\left| \frac{\langle \phi_m | \hat{H}_p \mid \phi_n \rangle}{E_n^{(0)} - E_m^{(0)}} \right| \ll 1 \qquad (n \neq m). \tag{9.24}$$

If the unperturbed energy levels $E_n^{(0)}$ and $E_m^{(0)}$ were equal (i.e., degenerate) then condition (9.24) would break down. Degenerate energy levels require an approach that is different from the nondegenerate treatment. This question will be taken up in the following section.

Example 9.1 (Charged oscillator in an electric field)

A particle of charge q and mass m, which is moving in a one-dimensional harmonic potential of frequency ω, is subject to a *weak* electric field $\mathcal{E}$ in the x-direction.

(a) Find the exact expression for the energy.

(b) Calculate the energy to first nonzero correction and compare it with the exact result obtained in (a).

Solution

The interaction between the oscillating charge and the external electric field gives rise to a term $\hat{H}_P = q\mathcal{E}\hat{X}$ that needs to be added to the Hamiltonian of the oscillator:

$$\hat{H} = \hat{H}_0 + \hat{H}_p = -\frac{\hbar^2}{2m}\frac{d^2}{dX^2} + \frac{1}{2}m\omega^2\hat{X}^2 + q\mathcal{E}\hat{X}. \tag{9.25}$$

(a) First, note that the eigenenergies of this Hamiltonian can be obtained exactly without resorting to any perturbative treatment. A variable change $\hat{y} = \hat{X} + q\mathcal{E}/(m\omega^2)$ leads to

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dy^2} + \frac{1}{2}m\omega^2\hat{y}^2 - \frac{q^2\mathcal{E}^2}{2m\omega^2}. \tag{9.26}$$

This is the Hamiltonian of a harmonic oscillator from which a constant, $q^2\mathcal{E}^2/(2m\omega^2)$, is subtracted. The exact eigenenergies can thus be easily inferred:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega - \frac{q^2\mathcal{E}^2}{2m\omega^2}. \tag{9.27}$$

This simple example allows us to compare the exact and approximate eigenenergies.

(b) Let us now turn to finding the approximate eigenvalues of $\hat{H}$ by means of perturbation theory. Since the electric field is weak, we can treat $\hat{H}_p$ as a perturbation.

Note that the first-order correction to the energy, $E_n^{(1)} = q\langle n \mid \hat{X} \mid n\rangle$, is zero (since $\langle n \mid \hat{X} \mid n\rangle = 0$), but the second-order correction is not:

$$E_n^{(2)} = q^2\mathcal{E}^2 \sum_{m\neq n} \frac{\left|\langle m \mid \hat{X} \mid n\rangle\right|^2}{E_n^{(0)} - E_m^{(0)}}. \tag{9.28}$$

Since $E_n^{(0)} = \left(n + \frac{1}{2}\right)\hbar\omega$, and using the relations

$$\langle n+1 \mid \hat{X} \mid n\rangle = \sqrt{n+1}\sqrt{\frac{\hbar}{2m\omega}}, \qquad \langle n-1 \mid \hat{X} \mid n\rangle = \sqrt{n}\sqrt{\frac{\hbar}{2m\omega}}, \tag{9.29}$$

$$E_n^{(0)} - E_{n-1}^{(0)} = \hbar\omega, \qquad E_n^{(0)} - E_{n+1}^{(0)} = -\hbar\omega, \tag{9.30}$$

we can reduce (9.28) to

$$\begin{aligned} E_n^{(2)} &= q^2\mathcal{E}^2\left[\frac{|\langle n+1 \mid \hat{X} \mid n\rangle|^2}{E_n^{(0)} - E_{n+1}^{(0)}} + \frac{|\langle n-1 \mid \hat{X} \mid n\rangle|^2}{E_n^{(0)} - E_{n-1}^{(0)}}\right] \\ &= -\frac{q^2\mathcal{E}^2}{2m\omega^2}; \end{aligned} \tag{9.31}$$

hence the energy is given to second order by

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} = \left(n + \frac{1}{2}\right)\hbar\omega - \frac{q^2\mathcal{E}^2}{2m\omega^2}. \tag{9.32}$$

This agrees fully with the exact energy found in (9.27).

Similarly, using (9.19) along with (9.29) and (9.30), we can easily ascertain that $\mid \psi_n^{(1)}\rangle$ is given by

$$\mid \psi_n^{(1)}\rangle = \frac{q\mathcal{E}}{\hbar\omega}\sqrt{\frac{\hbar}{2m\omega}}\left\{\sqrt{n} \mid n-1\rangle - \sqrt{n+1} \mid n+1\rangle\right\}; \tag{9.33}$$

hence the state $\mid \psi_n\rangle$ is given to first order by

$$\mid \psi_n\rangle = \mid n\rangle + \frac{q\mathcal{E}}{\hbar\omega}\sqrt{\frac{\hbar}{2m\omega}}\left\{\sqrt{n} \mid n-1\rangle - \sqrt{n+1} \mid n+1\rangle\right\}, \tag{9.34}$$

where $\mid n\rangle$ is the exact eigenstate of the nth excited state of a one-dimensional harmonic oscillator.

Example 9.2 (The Stark effect)

(a) Study the effect of an external uniform weak electric field, which is directed along the positive z-axis, $\vec{\mathcal{E}} = \mathcal{E}\hat{k}$, on the ground state of a hydrogen atom; ignore the spin degrees of freedom.

(b) Find an approximate value for the polarizability of the hydrogen atom.

Solution

(a) The effect that an external electric field has on the energy levels of an atom is called the Stark effect. In the absence of an electric field, the (unperturbed) Hamiltonian of the hydrogen atom (in CGS units) is:

$$\hat{H}_0 = \frac{\hat{\vec{p}}^2}{2\mu} - \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{r}. \tag{9.35}$$

The eigenfunctions of this Hamiltonian, $\psi_{nlm}(\vec{r})$, were obtained in Chapter 6; they are given by

$$\langle r\theta\varphi \mid nlm\rangle = \psi_{nlm}(r,\theta,\varphi) = R_{nl}(r)Y_{lm}(\theta,\varphi). \tag{9.36}$$

When the electric field is turned on, the interaction between the atom and the field generates a term $\hat{H}_p = e\vec{\mathcal{E}}\cdot\vec{r} = e\mathcal{E}\hat{Z}$ that needs to be added to $\hat{H}_0$.

Since the excited states of the hydrogen atom are degenerate while the ground state is not, nondegenerate perturbation theory applies only to the ground state, $\psi_{100}(\vec{r})$. Ignoring the spin degrees of freedom, the energy of this system to second-order perturbation is given as follows (see (9.23)):

$$E_{100} = E^{(0)}_{100} + e\mathcal{E}\langle 100 \mid \hat{Z} \mid 100\rangle + e^2\mathcal{E}^2 \sum_{nlm\neq 100} \frac{\left|\langle nlm \mid \hat{Z} \mid 100\rangle\right|^2}{E^{(0)}_{100} - E^{(0)}_{nlm}}. \tag{9.37}$$

The term

$$\langle 100 \mid \hat{Z} \mid 100\rangle = \int |\psi_{100}(\vec{r})|^2 z\, d^3r \tag{9.38}$$

is zero, since $\hat{Z}$ is odd under parity and $\psi_{100}(\vec{r})$ has a definite parity. This means that there can be no correction term to the energy which is proportional to the electric field and hence there is no linear Stark effect. The underlying physics behind this is that when the hydrogen atom is in its ground state, it has no permanent electric dipole moment. We are left then with only a *quadratic* dependence of the energy (9.37) on the electric field. This is called the quadratic Stark effect. This correction, which is known as the energy shift ΔE, is given by

$$\Delta E = e^2\mathcal{E}^2 \sum_{nlm\neq 100} \frac{\left|\langle nlm \mid \hat{Z} \mid 100\rangle\right|^2}{E^{(0)}_{100} - E^{(0)}_{nlm}}. \tag{9.39}$$

(b) Let us now estimate the value of the polarizability of the hydrogen atom. The *polarizability* α of an atom which is subjected to an electric field $\vec{\mathcal{E}}$ is given in terms of the energy shift ΔE as

$$\alpha = -2\frac{\Delta E}{\mathcal{E}^2}. \tag{9.40}$$

Substituting (9.39) into (9.40), we obtain the polarizability of the hydrogen atom in its ground state:

$$\alpha = -2e^2 \sum_{nlm\neq 100} \frac{\left|\langle nlm \mid \hat{Z} \mid 100\rangle\right|^2}{E^{(0)}_{100} - E^{(0)}_{nlm}}. \tag{9.41}$$

To estimate this sum, let us assume that the denominator is constant. Since $n \geq 2$, we can write

$$E^{(0)}_{100} - E^{(0)}_{nlm} \leq E_{100} - E_{200} = \frac{e^2}{2a_0}\left(-1 + \frac{1}{4}\right) = -\frac{3e^2}{8a_0}; \tag{9.42}$$

hence

$$\alpha \leq \frac{16a_0}{3} \sum_{nlm \neq 100} |\langle nlm \mid \hat{Z} \mid 100\rangle|^2, \tag{9.43}$$

where

$$\begin{aligned} \sum_{nlm \neq 100} |\langle nlm \mid \hat{Z} \mid 100\rangle|^2 &= \sum_{\text{all } nlm} |\langle nlm \mid \hat{Z} \mid 100\rangle|^2 \\ &= \langle 100 \mid \hat{Z} \left(\sum_{\text{all } nlm} \mid nlm\rangle\langle nlm \mid \right) \hat{Z} \mid 100\rangle \\ &= \langle 100 \mid \hat{Z}^2 \mid 100\rangle; \end{aligned} \tag{9.44}$$

in deriving this relation, we have used the facts that $\langle 100 \mid \hat{Z} \mid 100\rangle = 0$ and that the set of states $\mid nlm\rangle$ is complete. Now since $z = r\cos\theta$ and $\langle r\theta\varphi \mid 100\rangle = R_{10}(r)Y_{00}(\theta, \varphi) = R_{10}(r)/\sqrt{4\pi}$, we immediately obtain

$$\langle 100 \mid \hat{Z}^2 \mid 100\rangle = \frac{1}{4\pi} \int_0^\infty r^4 R_{10}^2(r)\, dr \int_0^\pi \sin\theta \cos^2\theta\, d\theta \int_0^{2\pi} d\varphi = a_0^2. \tag{9.45}$$

Substituting (9.45) and (9.44) into (9.43), we see that the polarizability for hydrogen has an upper limit

$$\alpha \leq \frac{16}{3} a_0^3. \tag{9.46}$$

This limit, which is obtained from perturbation theory, agrees with the exact value $\alpha = \frac{9}{2}a_0^3$.

9.2.2 Degenerate Perturbation Theory

In the discussion above, we have considered only systems with nondegenerate $\hat{H}_0$. We now apply perturbation theory to determine the energy spectrum and the states of a system whose unperturbed Hamiltonian $\hat{H}_0$ is degenerate:

$$\hat{H} \mid \psi_n\rangle = (\hat{H}_0 + \hat{H}_p) \mid \psi_n\rangle = E_n \mid \psi_n\rangle. \tag{9.47}$$

If, for instance, the level of energy $E_n^{(0)}$ is f-fold degenerate (i.e., there exists a set of f different eigenstates $\mid \phi_{n_\alpha}\rangle$, where $\alpha = 1, 2, \ldots, f$, that correspond to the same eigenenergy $E_n^{(0)}$), we have

$$\hat{H}_0 \mid \phi_{n_\alpha}\rangle = E_n^{(0)} \mid \phi_{n_\alpha}\rangle \qquad (\alpha = 1, 2, \ldots, f), \tag{9.48}$$

where α stands for one or more quantum numbers; the energy eigenvalues $E_n^{(0)}$ are independent of α.

In the zeroth-order approximation we can write the eigenfunction $\mid \psi_n\rangle$ as a linear combination in terms of $\mid \phi_{n_\alpha}\rangle$:

$$\mid \psi_n\rangle = \sum_{\alpha=1}^{f} a_\alpha \mid \phi_{n_\alpha}\rangle. \tag{9.49}$$

Considering the states $|\phi_{n_\alpha}\rangle$ to be orthonormal with respect to the label α (i.e., $\langle\phi_{n_\alpha}|\phi_{n_\beta}\rangle = \delta_{\alpha,\beta}$) and $|\psi_n\rangle$ to be normalized, $\langle\psi_n|\psi_n\rangle = 1$, we can ascertain that the coefficients a_α obey the relation

$$\langle\psi_n|\psi_n\rangle = \sum_{\alpha,\beta} a_\alpha^* a_\beta \delta_{\alpha,\beta} = \sum_{\alpha=1}^{f} |a_\alpha|^2 = 1. \tag{9.50}$$

In what follows we are going to show how to determine these coefficients and the first-order corrections to the energy. For this, let us substitute (9.48) and (9.49) into (9.47):

$$\sum_\alpha \left[E_n^{(0)}|\phi_{n_\alpha}\rangle + \hat{H}_p|\phi_{n_\alpha}\rangle\right] a_\alpha = E_n \sum_\alpha a_\alpha |\phi_{n_\alpha}\rangle. \tag{9.51}$$

The multiplication of both sides of this equation by $\langle\phi_{n_\beta}|$ leads to

$$\sum_\alpha a_\alpha \left[E_n^{(0)}\delta_{\alpha,\beta} + \langle\phi_{n_\beta}|\hat{H}_p|\phi_{n_\alpha}\rangle\right] = E_n \sum_\alpha a_\alpha \delta_{\alpha,\beta} \tag{9.52}$$

or to

$$a_\beta E_n = a_\beta E_n^{(0)} + \sum_{\alpha=1}^{f} a_\alpha \langle\phi_{n_\beta}|\hat{H}_p|\phi_{n_\alpha}\rangle, \tag{9.53}$$

where we have used $\langle\phi_{n_\beta}|\phi_{n_\alpha}\rangle = \delta_{\beta,\alpha}$. We can rewrite (9.53) as follows:

$$\sum_{\alpha=1}^{f} \left(\hat{H}_{p_{\beta\alpha}} - E_n^{(1)}\delta_{\alpha,\beta}\right) a_\alpha = 0 \qquad (\beta = 1, 2, \ldots, f), \tag{9.54}$$

with $\hat{H}_{p_{\beta\alpha}} = \langle\phi_{n_\beta}|\hat{H}_p|\phi_{n_\alpha}\rangle$ and $E_n^{(1)} = E_n - E_n^{(0)}$. This is a system of f homogeneous linear equations for the coefficients a_α. These coefficients are nonvanishing only when the determinant $|\hat{H}_{p_{\alpha\beta}} - E_n^{(1)}\delta_{\alpha,\beta}|$ is zero:

$$\begin{vmatrix} \hat{H}_{p_{11}} - E_n^{(1)} & \hat{H}_{p_{12}} & \hat{H}_{p_{13}} & \cdots & \hat{H}_{p_{1f}} \\ \hat{H}_{p_{21}} & \hat{H}_{p_{22}} - E_n^{(1)} & \hat{H}_{p_{23}} & \cdots & \hat{H}_{p_{2f}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{H}_{p_{f1}} & \hat{H}_{p_{f2}} & \hat{H}_{p_{f3}} & \cdots & \hat{H}_{p_{ff}} - E_n^{(1)} \end{vmatrix} = 0. \tag{9.55}$$

This is an fth degree equation in $E_n^{(1)}$ and in general it has f different real roots, $E_{n_\alpha}^{(1)}$. These roots are the first-order correction to the eigenvalues, E_{n_α}, of $\hat{H}$. To find the coefficients a_α, we need simply to substitute these roots into (9.54) and then solve the resulting expression. Knowing these coefficients, we can then determine the eigenfunctions, $|\psi_n\rangle$, of $\hat{H}$ in the zeroth approximation from (9.49).

The roots $E_{n_\alpha}^{(1)}$ of (9.55) are in general different. In this case the eigenvalues $\hat{H}$ are not degenerate, hence the f-fold degenerate level $E_n^{(0)}$ of the unperturbed problem is split into f different levels E_{n_α}: $E_{n_\alpha} = E_n^{(0)} + E_{n_\alpha}^{(1)}$, $\alpha = 1, 2, \ldots, f$. In this way, the perturbation lifts the degeneracy. The lifting of the degeneracy may be either total or partial, depending on whether all the roots of (9.55), or only some of them, are different.

In summary, to determine the eigenvalues to first-order and the eigenstates to zeroth order for an f-fold degenerate level from perturbation theory, we proceed as follows:

- First, for each f-fold degenerate level, determine the $f \times f$ matrix of the perturbation $\hat{H}_p$:

$$H_p = \begin{pmatrix} \hat{H}_{p_{11}} & \hat{H}_{p_{12}} & \cdots & \hat{H}_{p_{1f}} \\ \hat{H}_{p_{21}} & \hat{H}_{p_{22}} & \cdots & \hat{H}_{p_{2f}} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{H}_{p_{f1}} & \hat{H}_{p_{f2}} & \cdots & \hat{H}_{p_{ff}} \end{pmatrix}, \tag{9.56}$$

  where $\hat{H}_{p_{\alpha\beta}} = \langle \phi_{n_\alpha} | \hat{H}_p | \phi_{n_\beta} \rangle$.

- Second, diagonalize this matrix and find the f eigenvalues $E^{(1)}_{n_\alpha}$ ($\alpha = 1, 2, \ldots, f$) and their corresponding eigenvectors

$$a_\alpha = \begin{pmatrix} a_{\alpha_1} \\ a_{\alpha_2} \\ \vdots \\ a_{\alpha_f} \end{pmatrix} \qquad (\alpha = 1, 2, \ldots, f). \tag{9.57}$$

- Finally, the energy eigenvalues are given to first order by

$$E_{n_\alpha} = E^{(0)}_n + E^{(1)}_{n_\alpha} \qquad (\alpha = 1, 2, \ldots, f) \tag{9.58}$$

  and the corresponding eigenvectors are given to zero order by

$$|\psi_{n_\alpha}\rangle = \sum_{\beta=1}^{f} a_{\alpha\beta} |\phi_{n_\beta}\rangle. \tag{9.59}$$

Example 9.3 (The Stark effect of hydrogen)
Using first-order (degenerate) perturbation theory, calculate the energy levels of the $n = 2$ states of a hydrogen atom placed in an external uniform weak electric field along the positive z-axis.

Solution
In the absence of any external electric field, the first excited state (i.e., $n = 2$) is fourfold degenerate: the states $|\, nlm\rangle = |\, 200\rangle$, $|\, 210\rangle$, $|\, 211\rangle$, and $|\, 21-1\rangle$ have the same energy $E_2 = -R_y/4$, where $R_y = \mu e^4/(2\hbar^2) = 13.6\,\text{eV}$ is the Rydberg constant.

When the external electric field is turned on, some energy levels will split. The energy due to the interaction between the dipole moment of the electron ($\vec{d} = -e\vec{r}$) and the external electric field ($\vec{\mathcal{E}} = \mathcal{E}\vec{k}$) is given by

$$\hat{H}_p = -\vec{d} \cdot \vec{\mathcal{E}} = e\vec{r} \cdot \vec{\mathcal{E}} = e\mathcal{E}\hat{Z}. \tag{9.60}$$

To calculate the eigenenergies, we need to determine and then diagonalize the 4×4 matrix elements of $\hat{H}_p$: $\langle 2l'm' \mid \hat{H}_p \mid 2lm\rangle = e\mathcal{E}\langle 2l'm' \mid \hat{Z} \mid 2lm\rangle$. The matrix elements $\langle 2l'm' \mid \hat{Z} \mid 2lm\rangle$

can be calculated more simply by using the relevant selection rules and symmetries. First, since $\hat{Z}$ does not depend on the azimuthal angle φ, $z = r\cos\theta$, the elements $\langle 2l'm' \mid \hat{Z} \mid 2lm\rangle$ are nonzero only if $m' = m$. Second, as Z is odd, the states $\mid 2l'm'\rangle$ and $\mid 2lm\rangle$ must have opposite parities so that $\langle 2l'm' \mid \hat{Z} \mid 2lm\rangle$ does not vanish. Therefore, the only nonvanishing matrix elements are those that couple the 2s and 2p states (with $m = 0$); that is, between $\mid 200\rangle$ and $\mid 210\rangle$. In this case we have

$$\begin{aligned}\langle 200 \mid \hat{Z} \mid 210\rangle &= \int_0^\infty R_{20}^*(r)R_{21}(r)r^2dr \int Y_{00}^*(\Omega)zY_{10}(\Omega)\,d\Omega \\ &= \sqrt{\frac{4\pi}{3}}\int_0^\infty R_{20}(r)R_{21}(r)r^3dr \int Y_{00}^*(\Omega)Y_{10}^2(\Omega)\,d\Omega \\ &= -3a_0, \end{aligned} \tag{9.61}$$

since $z = r\cos\theta = \sqrt{4\pi/3}rY_{10}(\Omega)$, $\langle \vec{r} \mid 200\rangle = R_{20}(r)Y_{00}(\Omega)$, $\langle \vec{r} \mid 210\rangle = R_{21}(r)Y_{10}(\Omega)$, and $d\Omega = \sin\theta\, d\theta d\varphi$; $a_0 = \hbar^2/(\mu e^2)$ is the Bohr radius. Using the notations $\mid 1\rangle = \mid 200\rangle$, $\mid 2\rangle = \mid 211\rangle$, $\mid 3\rangle = \mid 210\rangle$, and $\mid 4\rangle = \mid 21-1\rangle$, we can write the matrix of H_p as

$$H_p = \begin{pmatrix} \langle 1 \mid \hat{H}_p \mid 1\rangle & \langle 1 \mid \hat{H}_p \mid 2\rangle & \langle 1 \mid \hat{H}_p \mid 3\rangle & \langle 1 \mid \hat{H}_p \mid 4\rangle \\ \langle 2 \mid \hat{H}_p \mid 1\rangle & \langle 2 \mid \hat{H}_p \mid 2\rangle & \langle 2 \mid \hat{H}_p \mid 3\rangle & \langle 2 \mid \hat{H}_p \mid 4\rangle \\ \langle 3 \mid \hat{H}_p \mid 1\rangle & \langle 3 \mid \hat{H}_p \mid 2\rangle & \langle 3 \mid \hat{H}_p \mid 3\rangle & \langle 3 \mid \hat{H}_p \mid 4\rangle \\ \langle 4 \mid \hat{H}_p \mid 1\rangle & \langle 4 \mid \hat{H}_p \mid 2\rangle & \langle 4 \mid \hat{H}_p \mid 3\rangle & \langle 4 \mid \hat{H}_p \mid 4\rangle \end{pmatrix} \tag{9.62}$$

or as

$$H_p = -3e\mathcal{E}a_0\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{9.63}$$

The diagonalization of this matrix leads to the following eigenvalues:

$$E_2^{(1)}{}_1 = -3e\mathcal{E}a_0, \qquad E_2^{(1)}{}_2 = E_2^{(1)}{}_3 = 0, \qquad E_2^{(1)}{}_4 = 3e\mathcal{E}a_0. \tag{9.64}$$

Thus, the energy levels of the $n = 2$ states are given to first order by

$$E_{2_1} = -\frac{R_y}{4} - 3e\mathcal{E}a_0, \quad E_{2_2} = E_{2_3} = -\frac{R_y}{4}, \quad E_{2_4} = -\frac{R_y}{4} + 3e\mathcal{E}a_0. \tag{9.65}$$

The corresponding eigenvectors to zeroth order are

$$\mid \psi_2\rangle_1 = \frac{1}{\sqrt{2}}(\mid 200\rangle + \mid 210\rangle), \qquad \mid \psi_2\rangle_2 = \mid 211\rangle, \tag{9.66}$$

$$\mid \psi_2\rangle_3 = \mid 21-1\rangle, \qquad \mid \psi_2\rangle_4 = \frac{1}{\sqrt{2}}(\mid 200\rangle - \mid 210\rangle). \tag{9.67}$$

This perturbation has only partially removed the degeneracy of the $n = 2$ level; the states $\mid 211\rangle$ and $\mid 21-1\rangle$ still have the same energy $E_3 = E_4 = -R_y/4$.

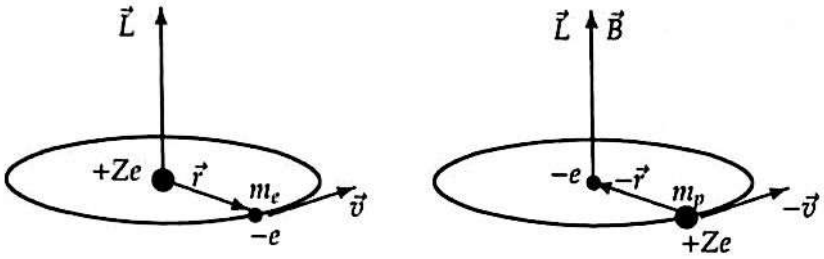


Figure 9.1: (Left) An electron moving in a circular orbit as seen by the nucleus. (Right) The same motion as seen by the electron within its rest frame; the electron sees the nucleus moving in a circular orbit around it.

9.2.3 Fine Structure and the Anomalous Zeeman Effect

One of the most useful applications of perturbation theory is to calculate the energy corrections for the hydrogen atom, notably the corrections due to the fine structure and the Zeeman effect. The fine structure is in turn due to two effects: spin–orbit coupling and the relativistic correction. Let us look at these corrections separately.

9.2.3.1 Spin–Orbit Coupling

The spin–orbit coupling in hydrogen arises from the interaction between the electron's spin magnetic moment, $\vec{\mu}_S = -e\vec{S}/(m_e c)$, and the proton's orbital magnetic field $\vec{B}$.

The origin of the magnetic field experienced by the electron moving at $\vec{v}$ in a circular orbit around the proton can be explained classically as follows. The electron, within its rest frame, sees the proton moving at $-\vec{v}$ in a circular orbit around it (Figure 9.1). From classical electrodynamics, the magnetic field experienced by the electron is

$$\vec{B} = -\frac{1}{c}\vec{v}\times\vec{E} = -\frac{1}{m_e c}\vec{p}\times\vec{E} = \frac{1}{m_e c}\vec{E}\times\vec{p}, \tag{9.68}$$

where $\vec{p} = m_e\vec{v}$ is the linear momentum of the electron and $\vec{E}$ is the electric field generated by the proton's Coulomb's field: $\vec{E}(\vec{r}) = (e/r^2)(\vec{r}/r) = e\vec{r}/r^3$. For a more general problem of hydrogen-like atoms—atoms with one valence electron outside a closed shell—where an electron moves in the (central) Coulomb potential of a nucleus $V(r) = -e\phi(r)$, the electric field is

$$\vec{E}(\vec{r}) = -\vec{\nabla}\phi(r) = \frac{1}{e}\vec{\nabla}V(r) = \frac{1}{e}\frac{\vec{r}}{r}\frac{dV}{dr}. \tag{9.69}$$

So the magnetic field of the nucleus calculated in the rest frame of the electron is obtained by inserting (9.69) into (9.68):

$$\vec{B} = \frac{1}{m_e c}\vec{E}\times\vec{p} = \frac{1}{em_e c}\frac{1}{r}\frac{dV}{dr}\vec{r}\times\vec{p} = \frac{1}{em_e c}\frac{1}{r}\frac{dV}{dr}\vec{L}, \tag{9.70}$$

where $\vec{L} = \vec{r}\times\vec{p}$ is the orbital angular momentum of the electron.

The interaction of the electron's spin dipole moment $\vec{\mu}_S$ with the orbital magnetic field $\vec{B}$ of the nucleus gives rise to the following interaction energy:

$$\hat{H}_{SO} = -\vec{\mu}_S \cdot \vec{B} = \frac{e}{m_e c}\vec{S}\cdot\vec{B} = \frac{1}{m_e^2 c^2}\frac{1}{r}\frac{dV}{dr}\vec{S}\cdot\vec{L}. \tag{9.71}$$

This energy turns out to be twice the observed spin–orbit interaction. This is due to the fact that (9.71) was calculated within the rest frame of the electron. This frame is not inertial, for the electron accelerates while moving in a circular orbit around the nucleus. For a correct treatment, we must transform to the rest frame of the nucleus (i.e., the lab frame). This transformation, which involves a relativistic transformation of velocities, gives rise to an additional motion resulting from the precession of $\vec{\mu}_S$; this is known as the *Thomas precession*. The precession of the electron's spin moment is a relativistic effect which occurs even in the absence of an external magnetic field. The transformation back to the rest frame of the nucleus leads to a reduction of the interaction energy (9.71) by a factor of 2:

$$\hat{H}_{SO} = \frac{1}{2m_e^2 c^2}\frac{1}{r}\frac{dV}{dr}\vec{S}\cdot\vec{L}. \tag{9.72}$$

As this relation was derived from a classical treatment, we can now obtain the corresponding quantum mechanical expression by replacing the dynamical variables with the corresponding operators:

$$\boxed{\hat{H}_{SO} = \frac{1}{2m_e^2 c^2}\frac{1}{r}\frac{d\hat{V}}{dr}\hat{\vec{S}}\cdot\hat{\vec{L}}.} \tag{9.73}$$

This is the *spin–orbit* energy. For a hydrogen's electron, $V(r) = -(e^2/(4\pi\varepsilon_0))/r$ and $dV/dr = (e^2/(4\pi\varepsilon_0))/r^2$, equation (9.73) reduces to

$$\hat{H}_{SO} = \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{2m_e^2 c^2}\frac{1}{r^3}\hat{\vec{S}}\cdot\hat{\vec{L}}. \tag{9.74}$$

We can now use perturbation theory to calculate the contribution of the spin–orbit interaction in a hydrogen atom:

$$\hat{H} = \frac{\hat{\vec{p}}^2}{2m_e} - \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{r} + \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{2m_e^2 c^2 r^3}\hat{\vec{S}}\cdot\hat{\vec{L}} = \hat{H}_0 + \hat{H}_{SO}, \tag{9.75}$$

where $\hat{H}_0$ is the unperturbed Hamiltonian and $\hat{H}_{SO}$ is the perturbation. To apply perturbation theory, we need to specify the unperturbed states—the eigenstates of $\hat{H}_0$. Since the spin of the hydrogen's electron is taken into account, the total wave function of $\hat{H}_0$ consists of a direct product of two parts: a spatial part and a spin part. To specify the eigenstates of $\hat{H}_0$, we have two choices: first, the joint eigenstates $|\, nlm_l m_s\rangle$ of $\hat{L}^2$, $\hat{S}^2$, $\hat{L}_z$, and $\hat{S}_z$ and, second, the joint eigenstates $|\, nljm\rangle$ of $\hat{L}^2$, $\hat{S}^2$, $\hat{J}^2$, and $\hat{J}_z$. While $\hat{H}_0$ is diagonal in both of these representations, $\hat{H}_{SO}$ is diagonal in the second but not in the first, because $\hat{H}_{SO}$ (or $\hat{\vec{S}}\cdot\hat{\vec{L}}$ to be precise) commutes with neither $\hat{L}_z$ nor with $\hat{S}_z$ (Chapter 7). Thus, if $\hat{H}_{SO}$ were included, the first choice would be a bad one, since we would be forced to diagonalize the matrix of $\hat{H}_{SO}$ within the states $|\, nlm_l m_s\rangle$; this exercise is nothing less than tedious and cumbersome. The second choice, however, is

ideal for our problem, since the first-order energy correction is given simply by the expectation value of the perturbation, because $\hat{H}_{SO}$ is already diagonal in this representation. We have shown in Chapter 7 (see Eq. (7.222))that the states $|\, nljm\rangle$,

$$\Psi_{n,l,j=l\pm\frac{1}{2},m} = R_{nl}(r)\left[\sqrt{\frac{l\mp m+\frac{1}{2}}{2l+1}}Y_{l,m+\frac{1}{2}}\left|\frac{1}{2},-\frac{1}{2}\right\rangle \pm \sqrt{\frac{l\pm m+\frac{1}{2}}{2l+1}}Y_{l,m-\frac{1}{2}}\left|\frac{1}{2},\frac{1}{2}\right\rangle\right], \tag{9.76}$$

are eigenstates of $\hat{\vec{S}}\cdot\hat{\vec{L}}$ and that the corresponding eigenvalues are given by

$$\langle nljm \,|\, \hat{\vec{L}}\cdot\hat{\vec{S}} \,|\, nljm\rangle = \frac{\hbar^2}{2}\left[j(j+1)-l(l+1)-\frac{3}{4}\right], \tag{9.77}$$

since $\hat{\vec{S}}\cdot\hat{\vec{L}} = \frac{1}{2}\left[\hat{J}^2-\hat{L}^2-\hat{S}^2\right]$.

The eigenvalues of (9.75) are then given to first-order correction by

$$E_{nlj} = E_n^{(0)} + \langle nljm_j \,|\, \hat{H}_{SO} \,|\, nljm_j\rangle = -\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{2a_0n^2} + E_{SO}^{(1)}, \tag{9.78}$$

where $E_n^{(0)} = -(e^2/(4\pi\varepsilon_0))/(2a_0n^2) = -(13.6/n^2)\,\text{eV}$ are the energy levels of hydrogen and $E_{SO}^{(1)}$ is the energy due to spin–orbit interaction:

$$E_{SO}^{(1)} = \langle nljm_j \,|\, \hat{H}_{SO} \,|\, nljm_j\rangle = \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{\hbar^2}{4m_e^2c^2}\left[j(j+1)-l(l+1)-\frac{3}{4}\right]\left\langle nl\left|\frac{1}{r^3}\right|nl\right\rangle. \tag{9.79}$$

Using the value of $\left\langle nl\left|r^{-3}\right|nl\right\rangle$ calculated in Chapter 6,

$$\left\langle nl\left|\frac{1}{r^3}\right|nl\right\rangle = \frac{2}{n^3l(l+1)(2l+1)a_0^3}, \tag{9.80}$$

we can rewrite (9.79) as

$$\begin{aligned} E_{SO}^{(1)} &= \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{\hbar^2}{2m_e^2c^2}\left[\frac{j(j+1)-l(l+1)-\frac{3}{4}}{n^3l(l+1)(2l+1)a_0^3}\right] \\ &= \left[\left(\frac{e^2}{4\pi\varepsilon_0}\right)\left(\frac{1}{2a_0n^2}\right)\right]\left(\frac{\hbar}{m_eca_0}\right)^2\frac{1}{n}\left[\frac{j(j+1)-l(l+1)-\frac{3}{4}}{l(l+1)(2l+1)}\right] \end{aligned} \tag{9.81}$$

or

$$\boxed{E_{SO}^{(1)} = \frac{|E_n^{(0)}|\alpha^2}{n}\left[\frac{j(j+1)-l(l+1)-\frac{3}{4}}{l(l+1)(2l+1)}\right]}, \tag{9.82}$$

where α is a dimensionless constant called the *fine structure constant*:

$$\alpha = \frac{\hbar}{m_eca_0} = \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{\hbar c} \simeq \frac{1}{137}. \tag{9.83}$$

Since $a_0 = ((4\pi\varepsilon_0)/e^2)\hbar^2/m_e$ and hence $E_n^{(0)} = -(e^2/(4\pi\varepsilon_0))/(2a_0n^2) = -\alpha^2 m_e c^2/(2n^2)$, we can express (9.82) in terms of α as

$$\boxed{E_{SO}^{(1)} = \frac{\alpha^4 m_e c^2}{2n^3}\left[\frac{j(j+1) - l(l+1) - \frac{3}{4}}{l(l+1)(2l+1)}\right].} \tag{9.84}$$

Note that when $l = 0$, the spin-orbit energy correction is zero; that is, when $l = 0$, the quantum number j becomes $j = 1/2$ and, hence, we infer from Eq. (9.79) that $E_{SO}^{(1)} = 0$ because $j(j+1) - l(l+1) - \frac{3}{4} = j(j+1) - \frac{3}{4} = \frac{3}{4} - \frac{3}{4} = 0$.

9.2.3.2 Relativistic Correction

Although the relativistic effect in hydrogen due to the motion of the electron is small, it can still be detected by spectroscopic techniques. The relativistic kinetic energy of the electron is given by $\hat{T} = \sqrt{\hat{p}^2c^2 + m_e^2c^4} - m_ec^2$, where m_ec^2 is the rest mass energy of the electron; an expansion of this relation to $\hat{p}^4$ yields

$$\sqrt{\hat{p}^2c^2 + m_e^2c^4} - m_ec^2 \simeq \frac{\hat{p}^2}{2m_e} - \frac{\hat{p}^4}{8m_e^3c^2} + \cdots. \tag{9.85}$$

When this term is included, the hydrogen's Hamiltonian becomes

$$\hat{H} = \frac{\hat{p}^2}{2m_e} - \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{r} - \frac{\hat{p}^4}{8m_e^3c^2} = \hat{H}_0 + \hat{H}_R, \tag{9.86}$$

where $\hat{H}_0 = \hat{p}^2/(2m_e) - e^2/r$ is the unperturbed Hamiltonian and

$$\hat{H}_R = -\frac{\hat{p}^4}{8m_e^3c^2}, \tag{9.87}$$

is the relativistic mass correction; this term results from the fact that the mass of the electron increases with speed as can be seen from the relativistic mass expression $m = m_0/\sqrt{1 - v^2/c^2}$ which shows that m increases as the speed v increases. A first-order perturbation treatment of this term yields:

$$E_R^{(1)} = \langle nljm_j \mid \hat{H}_R \mid nljm_j\rangle = -\frac{1}{8m_e^3c^2}\left\langle nljm_j\left|\hat{p}^4\right|nljm_j\right\rangle. \tag{9.88}$$

The value of $\left\langle nljm_j\left|\hat{p}^4\right|nljm_j\right\rangle$ was calculated in the last solved problem of Chapter 6 (see equation (6.359) on Page 426):

$$\left\langle nljm_j\left|\hat{p}^4\right|nljm_j\right\rangle = \frac{m_e^4e^8}{\hbar^4n^4}\left(\frac{8n}{2l+1} - 3\right) = \frac{\alpha^4m_e^4c^4}{n^4}\left(\frac{8n}{2l+1} - 3\right). \tag{9.89}$$

An insertion of this value in (9.88) leads to

$$\boxed{E_R^{(1)} = -\frac{\alpha^4 m_e c^2}{8n^4}\left(\frac{8n}{2l+1} - 3\right) = -\frac{\alpha^2|E_n^{(0)}|}{4n^2}\left(\frac{8n}{2l+1} - 3\right).} \tag{9.90}$$

Note that the spin–orbit and relativistic corrections (9.84) and (9.90) have the same order of magnitude, 10^{-3} eV, since $\alpha^2|E_n^{(0)}| \simeq 10^{-3}$ eV.

Remark

For a hydrogenlike atom having Z electrons, and if we neglect the spin–orbit interaction, we may use (9.90) to infer the atom's ground state energy:

$$\boxed{E_n = Z^2\left(E_n^{(0)} + E_R^{(1)}\right) = Z^2 E_n^{(0)}\left[1 + \frac{\alpha^2}{n}\left(\frac{2}{2l+1} - \frac{3}{4n}\right)\right]}, \tag{9.91}$$

where $E_n^{(0)} = -(e^2/(4\pi\varepsilon_0))/(2a_0n^2) = -\alpha^2 m_e c^2/(2n^2) = -(13.6/n^2)$ eV is the Bohr energy.

9.2.3.3 The Fine Structure of Hydrogen

The fine structure correction is obtained by adding the expressions for the spin–orbit and relativistic corrections (9.84) and (9.90):

$$E_{FS}^{(1)} = E_{SO}^{(1)} + E_R^{(1)} = \frac{\alpha^4 m_e c^2}{2n^3}\left[\frac{j(j+1) - l(l+1) - \frac{3}{4}}{l(l+1)(2l+1)}\right] - \frac{\alpha^4 m_e c^2}{8n^4}\left[\frac{8n}{2l+1} - 3\right], \tag{9.92}$$

where $j = l \pm \frac{1}{2}$. If $j = l + \frac{1}{2}$ a substitution of $l = j - \frac{1}{2}$ into (9.92) leads to

$$\begin{aligned} E_{FS}^{(1)} &= \frac{\alpha^4 m_e c^2}{8n^4}\left[\frac{4nj(j+1) - 4n\left(j-\frac{1}{2}\right)\left(j+\frac{1}{2}\right) - 3n}{\left(j-\frac{1}{2}\right)\left(j+\frac{1}{2}\right)(2j-1+1)} - \frac{8n}{2j-1+1} + 3\right] \\ &= \frac{\alpha^4 m_e c^2}{8n^4}\left[\frac{4nj-2n}{2j\left(j-\frac{1}{2}\right)\left(j+\frac{1}{2}\right)} - \frac{4n}{j} + 3\right] = \frac{\alpha^4 m_e c^2}{8n^4}\left[\frac{2n}{j\left(j+\frac{1}{2}\right)} - \frac{4n}{j} + 3\right] \\ &= \frac{\alpha^4 m_e c^2}{8n^4}\left[3 - \frac{4n}{j+\frac{1}{2}}\right]. \end{aligned} \tag{9.93}$$

Similarly, if $j = l - \frac{1}{2}$, and hence $l = j + \frac{1}{2}$, we can reduce (9.92) to

$$\begin{aligned} E_{FS}^{(1)} &= \frac{\alpha^4 m_e c^2}{8n^4}\left[\frac{4nj(j+1) - 4n\left(j+\frac{1}{2}\right)\left(j+\frac{3}{2}\right) - 3n}{\left(j+\frac{1}{2}\right)\left(j+\frac{3}{2}\right)(2j+1+1)} - \frac{8n}{2j+1+1} + 3\right] \\ &= \frac{\alpha^4 m_e c^2}{8n^4}\left[\frac{-4nj-6n}{2\left(j+\frac{1}{2}\right)\left(j+\frac{3}{2}\right)(j+1)} - \frac{4n}{j+1} + 3\right] \\ &= \frac{\alpha^4 m_e c^2}{8n^4}\left[\frac{-2n}{\left(j+\frac{1}{2}\right)(j+1)} - \frac{4n}{j+1} + 3\right] \\ &= \frac{\alpha^4 m_e c^2}{8n^4}\left[3 - \frac{4n}{j+\frac{1}{2}}\right]. \end{aligned} \tag{9.94}$$

As equations (9.93) and (9.94) show, the expressions for the fine structure correction corresponding to $j = l + \frac{1}{2}$ and $j = l - \frac{1}{2}$ are the same:

$$\boxed{E_{FS}^{(1)} = E_{SO}^{(1)} + E_R^{(1)} = \frac{\alpha^4 m_e c^2}{8n^4}\left(3 - \frac{4n}{j+\frac{1}{2}}\right) = \frac{\alpha^2 E_n^{(0)}}{4n^2}\left(\frac{4n}{j+\frac{1}{2}} - 3\right),} \tag{9.95}$$

where $E_n^{(0)} = -\alpha^2 m_e c^2/(2n^2)$ and $j = l \pm \frac{1}{2}$.

In sum, when including the fine structure, the hydrogen's Hamiltonian is given by

$$\hat{H} = \hat{H}_0 + \hat{H}_{FS} = \hat{H}_0 + \left(\hat{H}_{SO} + \hat{H}_R\right) = \frac{\hat{p}^2}{2m_e} - \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{r} + \left[\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{2m_e^2 c^2 r^3}\hat{\vec{S}}\cdot\hat{\vec{L}} - \frac{\hat{p}^4}{8m_e^3 c^2}\right]. \tag{9.96}$$

A first-order perturbation calculation of the energy levels of hydrogen, when including the fine structure, yields

$$\boxed{E_{nj} = E_n^{(0)} + E_{FS}^{(1)} = E_n^{(0)}\left[1 + \frac{\alpha^2}{4n^2}\left(\frac{4n}{j+\frac{1}{2}} - 3\right)\right],} \tag{9.97}$$

where $E_n^{(0)} = -(13.6/n^2)\,\text{eV}$. Unlike $E_n^{(0)}$, which is degenerate in l, each energy level E_{nj} is split into two levels $E_{n\,(l\pm\frac{1}{2})}$, since for a given value of l there are two values of j: $j = l \pm \frac{1}{2}$.

In addition to the fine structure, there is still another (smaller) effect which is known as the *hyperfine structure*. The hydrogen's hyperfine structure results from the interaction of the spin of the electron with the spin of the nucleus. When the hyperfine corrections are included, they would split each of the fine structure levels into a series of hyperfine levels. For instance, when the hyperfine coupling is taken into account in the ground state of hydrogen, it would split the $1S_{1/2}$ level into two hyperfine levels separated by an energy of $5.89 \times 10^{-6}\,\text{eV}$. This corresponds, when the atom makes a spontaneous transition from the higher hyperfine level to the lower one, to a radiation of $1.42 \times 10^9\,\text{Hz}$ frequency and 21 cm wavelength. We should note that most of the information we possess about interstellar hydrogen clouds had its origin in the radioastronomy study of this 21 cm line.

9.2.3.4 The Darwin Term

Another pertinent relativistic correction to the hydrogen's energy spectrum is provided by the *Darwin term*,

$$\hat{H}_{Darwin} = \frac{e^2\hbar^2}{8m_e^2 c^2 \varepsilon_0}\delta(\vec{r}), \tag{9.98}$$

which, as we will see in Chapter 12 (see Eq. 12.486), comes out naturally from the Dirac equation. This term originates from the fact that, in relativity, the position of the electron in an atom cannot be determined with infinite precision due to the fuzzy nature of its "trajectory." That is, the electron is not a point particle; it rather spreads over a "volume" of the size of its Compton wavelength $\lambda_c \sim \hbar/(m_e c)$ and, hence, the electron cannot be localized to an accuracy better than λ_c.

Due to $\delta(\vec{r})$ in (9.98), a first-order perturbation calculation shows that the contribution of the Darwin term to the total energy is non-zero *only* at the *origin*:

$$E_D^{(1)} = \langle nlm \mid \hat{H}_{Darwin} \mid nlm \rangle = \frac{e^2\hbar^2}{8m_e^2c^2\varepsilon_0}\langle nlm \mid \delta(\vec{r}) \mid nlm \rangle = \frac{e^2\hbar^2}{8m_e^2c^2\varepsilon_0}\left|\psi_{nlm}(0)\right|^2, \tag{9.99}$$

where $\langle \vec{r} \mid nlm \rangle = \psi_{nlm}(\vec{r}) \equiv R_{nl}(r)Y_{lm}(\theta,\varphi)$ as shown in Eq. (6.170).

Let us now focus on finding the wave function at the origin, $\psi_{nlm}(0)$. The radial function has a distinct property: at the origin, $R_{nl}(0)$ is not zero only for $l = 0$; it is zero for all $l \geq 1$ values. This result can be easily inferred from the properties of the associated Laguerre polynomial L_{n+l}^{2l+1} and the structure of $R_{nl}(r)$ as can be seen from Eq. (6.170) and Table 6.6 of Chapter 6; namely, since $R_{nl}(r) \propto (2r/(na_0))^l e^{-r/(na_0)} L_{n+l}^{2l+1}(2r/(na_0))$, the radial function is zero at the origin, $R_{nl}(0) = 0$, for $l \neq 0$ and not zero only for $l = 0$. For instance, the radial functions of the first few $S-$states are given by:

$$R_{10}(0) = \frac{2}{\sqrt{a_0^3}} \equiv \frac{2}{\sqrt{1^3 \times a_0^3}}, \qquad R_{20}(0) = \frac{1}{\sqrt{2a_0^3}} \equiv \frac{2}{\sqrt{2^3 \times a_0^3}}, \tag{9.100}$$

$$R_{30}(0) = \frac{2}{\sqrt{27a_0^3}} \equiv \frac{2}{\sqrt{3^3 \times a_0^3}}, \qquad R_{40}(0) = \frac{1}{\sqrt{16a_0^3}} \equiv \frac{2}{\sqrt{4^3 \times a_0^3}}, \cdots. \tag{9.101}$$

So, we can quite easily infer the general expression of the radial functions at the origin:

$$R_{nl}(0) = \begin{cases} \frac{2}{\sqrt{n^3a_0^3}} & l = 0 \\ 0 & l \geq 1 \end{cases} \tag{9.102}$$

where $a_0 = \left(\frac{4\pi\varepsilon_0}{e^2}\right)\frac{\hbar^2}{m_e}$ is the Bohr radius. Using (9.102) along with the relation $Y_{00} = 1/\sqrt{4\pi}$, we can obtain the wave function at the origin as follows:

$$\psi_{nlm}(0) = \begin{cases} R_{n0}(0)Y_{00}(0,\,0) = \frac{1}{\sqrt{\pi n^3a_0^3}} & l = 0 \\ 0 & l \geq 1 \end{cases} \tag{9.103}$$

This relation shows that the wave function is zero at the origin for the $p,\ d,\ f,\ \cdots$ orbitals (i.e., for $l = 1,\ 2, 3, \cdots$) and non-zero only for the $s-$states, $l = 0$.

Finally, inserting (9.103) into (9.99) , we see that the contribution of the Darwin term is not zero only for the $S-$statese, $l = 0$:

$$\begin{aligned} E_D^{(1)} &= \langle n00 \mid \hat{H}_{Darwin} \mid n00 \rangle = \frac{e^2\hbar^2}{8m_e^2c^2\varepsilon_0}\langle n00 \mid \delta(\vec{r}) \mid n00 \rangle = \frac{e^2\hbar^2}{8m_e^2c^2\varepsilon_0}\frac{1}{\pi n^3a_0^3} \\ &= \left[\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{2n^2a_0}\right]\frac{\hbar^2}{na_0^2m_e^2c^2} \equiv \left[\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{2n^2a_0}\right]\left(\frac{\hbar^2}{m_e^2c^2}\right)\left[\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{m_e}{\hbar^2}\right]^2\frac{1}{n} \\ &= \left[\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{2n^2a_0}\right]\left[\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{\hbar c}\right]^2\frac{1}{n} \\ &= -\frac{\alpha^2}{n}E_n^{(0)} \end{aligned} \tag{9.104}$$

where $E_n^{(0)} = -\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{2n^2a_0}$ and $\alpha = \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{\hbar c}$. Note that, since $E_n^{(0)}$ is also given by $E_n^{(0)} = \frac{1}{2n^2}m_e c^2\alpha^2$, we can write the Darwin contribution in the following three equivalent forms

$$\boxed{E_D^{(1)} = -\frac{\alpha^2}{n}E_n^{(0)} \equiv \frac{2n}{m_e c^2}\left(E_n^{(0)}\right)^2 \equiv \frac{1}{2n^3}m_e^2c^2\alpha^4, \qquad n = 1, 2, 3, \cdots.} \tag{9.105}$$

For instance, the contributions of the Darwin term to the $1S$, $2S$, and $3S$ states are given by

$$E_{D_{1S}}^{(1)} = \frac{1}{2}m_e^2c^2\alpha^4, \qquad E_{D_{2S}}^{(1)} = \frac{1}{16}m_e^2c^2\alpha^4, \qquad E_{D_{3S}}^{(1)} = \frac{1}{54}m_e^2c^2\alpha^4. \tag{9.106}$$

From (9.105), we see that the ratio of the contribution of the Darwin term to the energy of the hydrogen atom is of the order of α^2:

$$\frac{E_D^{(1)}}{|E_n^{(0)}|} = \frac{\alpha^2}{n}. \tag{9.107}$$

So, since $E_D^{(1)}$ is not zero only for $l = 0$, the Darwin term has no contribution to the spin-orbit coupling; as Eq. (9.79) shows, when $l = 0$, the spin-orbit energy correction is zero: $E_{SO}^{(1)} = 0$.

Remarks

- We should note that *fine structure is a relativistic effect*. In this chapter, we have introduced it in a rather heuristic manner by simply adding it to the Hamiltonian and treating it as a perturbation. However, and as we will show in Chapter 12 (see Subsection 12.4.20.2), the fine structure term appears naturally in the Dirac equation without the need to resort to any ad hoc procedure.
- Since the bracket-terms in (9.82), (9.90), and (9.95), are of the order of unity, the ratios of the spin–orbit, relativistic, fine structure, and Darwin term corrections to the energy of the hydrogen atom are of the order of α^2:

$$\frac{E_{SO}^{(1)}}{|E_n^{(0)}|} \simeq \frac{\alpha^2}{n}, \qquad \left|\frac{E_R^{(1)}}{E_n^{(0)}}\right| \simeq \frac{\alpha^2}{n}, \qquad \frac{E_{FS}^{(1)}}{|E_n^{(0)}|} \simeq \frac{\alpha^2}{n}, \qquad \frac{E_D^{(1)}}{|E_n^{(0)}|} = \frac{\alpha^2}{n}. \tag{9.108}$$

  For $n = 1$, all these terms are of the order of 10^{-4} since $\alpha^2 = (1/137)^2 \simeq 10^{-4}$.

9.2.3.5 The Anomalous Zeeman Effect

We now consider a hydrogen atom that is placed in an external uniform magnetic field $\vec{B}$. The effect of an external magnetic field on the atom is to cause a shift of its energy levels; this is called the *Zeeman effect*. In Chapter 6 we studied the Zeeman effect, but with one major omission: we ignored the spin of the electron. In this section we are going to take it into account. The interaction of the magnetic field with the electron's orbital and spin magnetic dipole moments, $\vec{\mu}_L$ and $\vec{\mu}_S$, gives rise to two energy terms, $-\vec{\mu}_L \cdot \vec{B}$ and $-\vec{\mu}_S \cdot \vec{B}$, whose sum we call the Zeeman energy:

$$\hat{H}_Z = -\vec{\mu}_L \cdot \vec{B} - \vec{\mu}_S \cdot \vec{B} = \frac{e}{2m_e c}\vec{L} \cdot \vec{B} + \frac{e}{m_e c}\vec{S} \cdot \vec{B} = \frac{e}{2m_e c}\left(\vec{L} + 2\vec{S}\right) \cdot \vec{B} = \frac{eB}{2m_e c}\left(\hat{L}_z + 2\hat{S}_z\right), \tag{9.109}$$

with $\vec{\mu}_L = -e\vec{L}/(2m_ec)$ and $\vec{\mu}_S = -\vec{S}/(m_ec)$; for simplicity, we have taken $\vec{B}$ along the z-axis: $\vec{B} = B\hat{z}$.

When a hydrogen atom is placed in an external magnetic field, its Hamiltonian is given by

$$\hat{H} = \hat{H}_0 + \hat{H}_{FS} + \hat{H}_Z. \tag{9.110}$$

Like $\hat{H}_{FS}$, the correction due to $\hat{H}_Z$ of (9.110) is expected to be small compared to $\hat{H}_0$; hence it can be treated perturbatively. We may now consider separately the cases where the magnetic field $\vec{B}$ is strong or weak. Strong or weak compared to what? Since $\hat{H}_{SO}$ and $\hat{H}_Z$ can be written as $\hat{H}_{SO} = W\vec{L}\cdot\vec{S}$ (see Eq. (9.74)) and since $\hat{H}_Z = B\mu_B(\hat{L}_z + 2\hat{S}_z)/\hbar$, we have $\hat{H}_Z/\hat{H}_{SO} \sim B\mu_B/W$, where μ_B is the Bohr magneton, $\mu_B = e\hbar/(2m_e)$. Thus, the cases $B \ll W/\mu_B$ and $B \gg W/\mu_B$ would correspond to the weak and strong magnetic fields, respectively.

The strong-field Zeeman effect

The effect of a strong external magnetic field on the hydrogen atom is called the *Paschen–Back effect*. If $\vec{B}$ is strong, $B \gg W/\mu_B$, the term $eB(\hat{L}_z + 2\hat{S}_z)/(2m_ec)$ will be much greater than the fine structure. Neglecting $\hat{H}_{FS}$, we can reduce (9.110) to

$$\hat{H} = \hat{H}_0 + \hat{H}_Z = \hat{H}_0 + \frac{eB}{2m_ec}\left(\hat{L}_z + 2\hat{S}_z\right). \tag{9.111}$$

Since $\hat{H}$ commutes with $\hat{H}_0$ (because $\hat{H}_0$ commutes with $\hat{L}_z$ and $\hat{S}_z$), they can be diagonalized by a common set of states, $|\, nlm_lm_s\rangle$:

$$\hat{H}\,|\, nlm_lm_s\rangle = \left[\hat{H}_0 + \frac{eB}{2m_ec}\left(\hat{L}_z + 2\hat{S}_z\right)\right]|\, nlm_lm_s\rangle = E_{nlm_lm_s}\,|\, nlm_lm_s\rangle, \tag{9.112}$$

where

$$E_{nlm_lm_s} = E_n^{(0)} + \frac{eB\hbar}{2m_ec}(m_l + 2m_s) = -\frac{e^2}{2a_0n^2} + \frac{eB\hbar}{2m_ec}(m_l + 2m_s). \tag{9.113}$$

The energy levels $E_n^{(0)}$ are thus shifted by an amount equal to $\Delta E = B\mu_B(m_l + 2m_s)$ with $\mu_B = e\hbar/(2m_ec)$, known as the Paschen–Back shift (Figure 9.2). When $\vec{B} = 0$ the degeneracy of each level of hydrogen is given by $g_n = 2\sum_{l=0}^{n-1}(2l+1) = 2n^2$; when $\vec{B} \neq 0$ states with the same value of $(m_l + 2m_s)$ are still degenerate.

The weak-field Zeeman effect

If $\vec{B}$ is weak, $B \ll W/\mu_B$, we need to consider all the terms in the Hamiltonian (9.110); the fine structure term $\hat{H}_{FS}$ will be the dominant perturbation. In the case where the Hamiltonian contains several perturbations at once, we should treat them individually starting with the most dominant, then the next, and so on. In this case the eigenstate should be selected to be one that diagonalizes the unperturbed Hamiltonian and the dominant perturbation[1]. In the weak-field Zeeman effect, since $\hat{H}_{FS}$ is the dominant perturbation, the best eigenstates to use are $|\, nljm_j\rangle$, for they simultaneously diagonalize $\hat{H}_0$ and $\hat{H}_{FS}$. Writing $\hat{L}_z + 2\hat{S}_z$ as $\hat{J}_z + \hat{S}_z$, where $\vec{\hat{J}} = \vec{\hat{L}} + \vec{\hat{S}}$ represents the total angular momentum of the electron, we may rewrite (9.110) as

$$\hat{H} = \hat{H}_0 + \hat{H}_{FS} + \hat{H}_Z = \hat{H}_0 + \hat{H}_{FS} + \frac{eB}{2m_ec}\left(\hat{J}_z + \hat{S}_z\right). \tag{9.114}$$

[1] When the various perturbations are of approximately equal size, a state that is a joint eigenstate of $\hat{H}_0$ and any perturbation would be an acceptable choice.

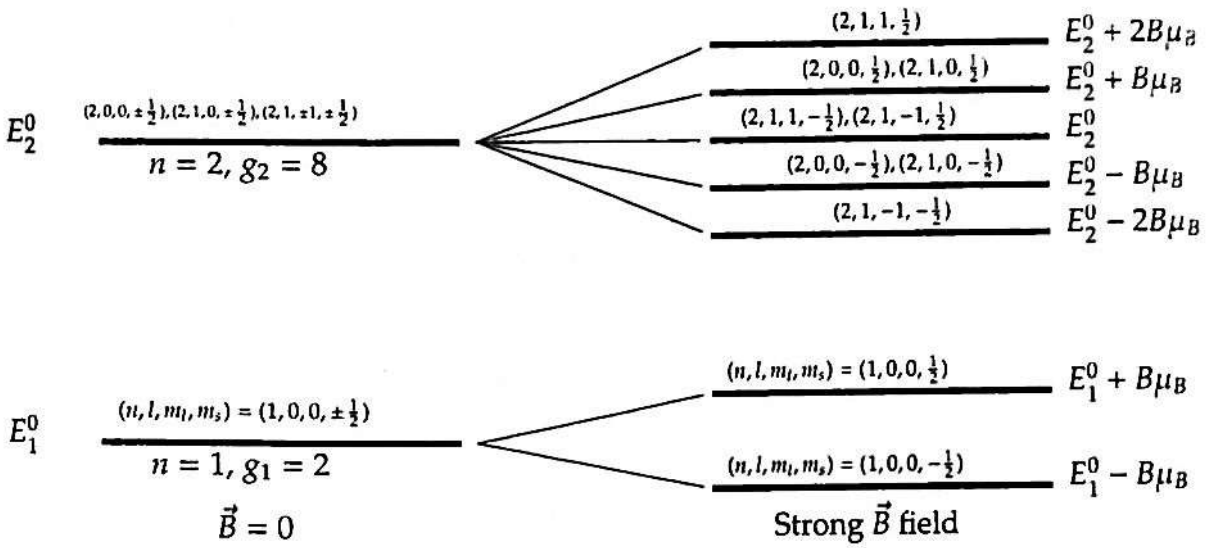


Figure 9.2: Splittings of the energy levels $n = 1$ and $n = 2$ of a hydrogen atom when placed in a *strong* external magnetic field; $\mu_B = e\hbar/(2m_e)$.

In a first-order perturbation calculation, the contribution of $\hat{H}_Z$ is given by

$$E_Z^{(1)} = \langle nljm_j \mid \hat{H}_Z \mid nljm_j \rangle = \frac{eB}{2m_e c}\langle nljm_j \mid \hat{J}_z + \hat{S}_z \mid nljm_j \rangle. \tag{9.115}$$

Since $\langle nljm_j \mid \hat{J}_z \mid nljm_j \rangle = \hbar m_j$ and using the expression of $\langle nljm_j \mid \hat{S}_z \mid nljm_j \rangle$ that was calculated in Chapter 7,

$$\begin{aligned}\langle nljm_j \mid \hat{S}_z \mid nljm_j \rangle &= \frac{\langle nljm_j \mid \hat{\vec{J}} \cdot \hat{\vec{S}} \mid nljm_j \rangle}{\hbar^2 j(j+1)}\langle nljm_j \mid \hat{J}_z \mid nljm_j \rangle \\ &= \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)}\hbar m_j,\end{aligned} \tag{9.116}$$

we can reduce (9.115) to

$$\boxed{E_Z^{(1)} = \frac{eB\hbar}{2m_e c}\left[1 + \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)}\right]m_j = \frac{eB\hbar}{2m_e c}g_j m_j = B\mu_B m_j g_j,} \tag{9.117}$$

where $\mu_B = e\hbar/(2m_e)$ is the Bohr magneton for the electron and g_j is the Landé factor or the gyromagnetic ratio:

$$g_j = 1 + \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)}. \tag{9.118}$$

This shows that when $l = 0$ and $j = s$ we have $g_s = 2$ and when $s = 0$ and $j = l$ we have $g_l = 1$. For instance, for an atomic state[2] such as $^2P_{3/2}$, (9.118) shows that its factor is given by

[2]We use here the spectroscopic notation where $^{(2s+1)}L_j$ designates an atomic state whose spin is s, its total angular momentum is j, and whose orbital angular momentum is L where the values $L = 0, 1, 2, 3, 4, 5, \ldots$ are designated, respectively, by the capital letters S, P, D, F, G, H, ... (see Chapter 8).

$g_{j=3/2} = \frac{4}{3}$, since $j = l + s = 1 + \frac{1}{2} = \frac{3}{2}$; this is how we infer the factor of any state:

$$\begin{array}{cccccccc} \text{State} & {}^2S_{1/2} & {}^2P_{1/2} & {}^2P_{3/2} & {}^2D_{3/2} & {}^2D_{5/2} & {}^2F_{5/2} & {}^2F_{7/2} \\ g_j & 2 & \frac{2}{3} & \frac{4}{3} & \frac{4}{5} & \frac{6}{5} & \frac{6}{7} & \frac{8}{7} \end{array} \tag{9.119}$$

From (9.118), we see that the Landé factors corresponding to the same l but different values of j (due to spin) are not equal, since for $s = \frac{1}{2}$ and $j = l \pm \frac{1}{2}$ we have

$$g_{j=l\pm\frac{1}{2}} = 1 \pm \frac{1}{2l+1} = \begin{cases} \frac{2l+2}{2l+1} & \text{for} \quad j = l + \frac{1}{2}, \\ \frac{2l}{2l+1} & \text{for} \quad j = l - \frac{1}{2}. \end{cases} \tag{9.120}$$

Combining (9.97), (9.114), and (9.117), we can write the energy of a hydrogen atom in a weak external magnetic field as follows:

$$\boxed{E_{nj} = E_n^{(0)} + E_{FS}^{(1)} + E_Z^{(1)} = E_n^{(0)} + \frac{\alpha^2 E_n^{(0)}}{4n^2}\left(\frac{4n}{j+\frac{1}{2}} - 3\right) + \frac{eB\hbar}{2m_e c} m_j g_j.} \tag{9.121}$$

The effect of the magnetic field on the atom is thus to split the energy levels with a spacing $\Delta E = B\mu_B m_j g_j$. Unlike the energy levels obtained in Chapter 6, where we ignored the electron's spin, the energy levels (9.121) are not degenerate in l. Each energy level j is split into an *even number* of $(2j + 1)$ sublevels corresponding to the $(2j + 1)$ values of m_j: $m_j = -j, -j + 1, \ldots, j - 1, j$. As displayed in Figure 9.3, the splittings between the sublevels corresponding to the same j are constant: the spacings between the sublevels corresponding to $j = l - 1/2$ are all equal to $\Delta\epsilon_1 = B\mu_B(2l)/(2l + 1)$, and the spacings between the $j = l + 1/2$ sublevels are equal to $\Delta\epsilon_2 = B\mu_B(2l + 2)/(2l + 1)$. In contrast to the normal Zeeman effect, however, the spacings between the split levels of the same l (and different values of j) are no longer constant, $\Delta\epsilon_1 \neq \Delta\epsilon_2$, since they depend on the Landé factor g_j; for a given value of j, there are two different values of g_j corresponding to $l = j \pm \frac{1}{2}$: $g_{j=l+1/2} = (2l + 2)/(2l + 1)$ and $g_{j=l-1/2} = (2l)/(2l + 1)$; see (9.120). This *unequal* spacing between the split levels is called the *anomalous* Zeeman effect.

9.3 The Variational Method

There exist systems whose Hamiltonians are known, but they cannot be solved exactly or by a perturbative treatment. That is, there is no closely related Hamiltonian that can be solved exactly or approximately by perturbation theory because the first order is not sufficiently accurate. One of the approximation methods that is suitable for solving such problems is the *variational method*, which is also called the Rayleigh–Ritz method. This method does not require knowledge of simpler Hamiltonians that can be solved exactly. The variational method is useful for determining upper bound values for the eigenenergies of a system whose Hamiltonian is known whereas its eigenvalues and eigenstates are not known. It is particularly useful for determining the ground state. It becomes quite cumbersome to determine the energy levels of the excited states.

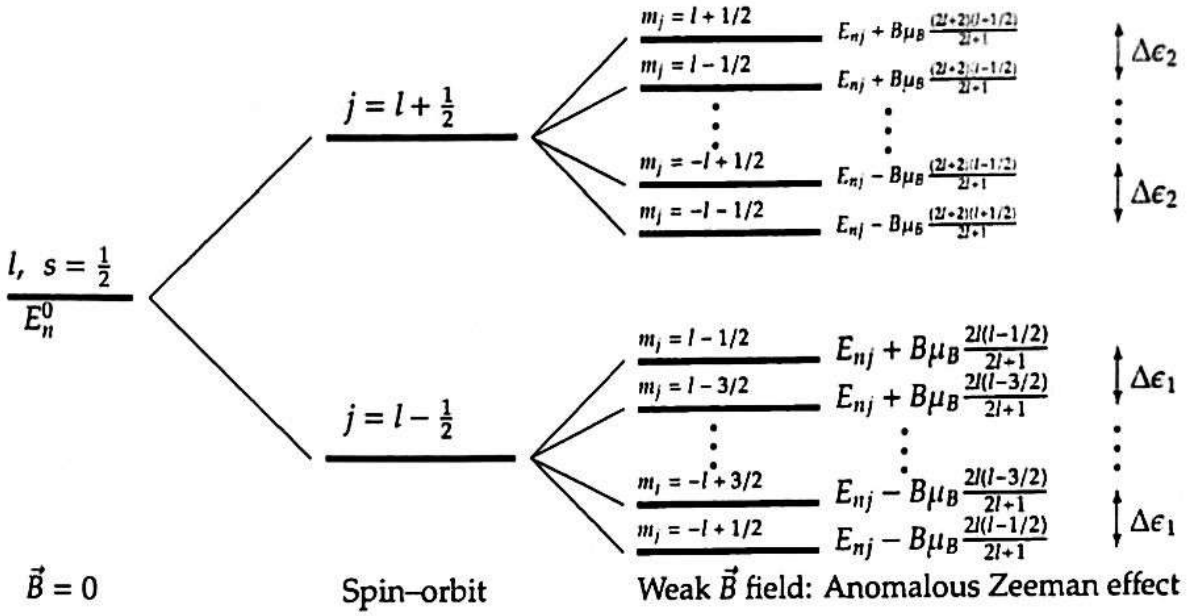


Figure 9.3: Splittings of a level l due to the spin–orbit interaction and to a weak external magnetic field, with $E_{nj} = E_n^0 + E_{SO}^{(1)}$. All the lower sublevels are equally spaced, $\Delta\epsilon_1 = B\mu_B(2l)/(2l+1)$, and so are the upper sublevels, $\Delta\epsilon_2 = B\mu_B(2l+2)/(2l+1)$, with $\mu_B = e\hbar/(2m_e)$.

In the context of the variational method, one does not attempt to solve the eigenvalue problem

$$\hat{H} \mid \psi\rangle = E \mid \psi\rangle, \tag{9.122}$$

but rather one uses a variational scheme to find the approximate eigenenergies and eigenfunctions from the variational equation

$$\delta E(\psi) = 0, \tag{9.123}$$

where $E(\psi)$ is the expectation value of the energy in the state $\mid \psi\rangle$:

$$E(\psi) = \frac{\langle \psi \mid \hat{H} \mid \psi\rangle}{\langle \psi \mid \psi\rangle}. \tag{9.124}$$

If $\mid \psi\rangle$ depends on a parameter α, $E(\psi)$ will also depend on α. The variational ansatz (9.123) enables us to vary α so as to minimize $E(\psi)$. The minimum value of $E(\psi)$ provides an upper limit approximation for the true energy of the system.

The variational method is particularly useful for determining the ground state energy and its eigenstate without explicitly solving the Schrödinger equation. Note that for any (arbitrary) trial function $\mid \psi\rangle$ we choose, the energy E as given by (9.124) is always larger than the exact energy E_0:

$$E = \frac{\langle \psi \mid H \mid \psi\rangle}{\langle \psi \mid \psi\rangle} \geq E_0; \tag{9.125}$$

the equality condition occurs only when $\mid \psi\rangle$ is proportional to the true ground state $|\psi_0\rangle$. To prove this, we simply expand the trial function $\mid \psi\rangle$ in terms of the exact eigenstates of $\hat{H}$:

$$\mid \psi\rangle = \sum_n a_n \mid \phi_n\rangle, \tag{9.126}$$

with

$$H \mid \phi_n\rangle = E_n \mid \phi_n\rangle, \tag{9.127}$$

and since $E_0 \geq E_n$ for nondegenerate one-dimensional bound systems, we have

$$E = \frac{\langle\psi \mid H|\psi\rangle}{\langle\psi \mid \psi\rangle} = \frac{\sum_n |a_n|^2 E_n}{\sum_n |a_n|^2} \geq \frac{E_0 \sum_n |a_n|^2}{\sum_n |a_n|^2} = E_0, \tag{9.128}$$

which proves (9.125).

To calculate the ground state energy, we need to carry out the following four steps:

- First, based on physical intuition, make an educated guess of a trial function that takes into account all the physical properties of the ground state (symmetries, number of nodes, smoothness, behavior at infinity, etc.). For the properties you are not sure about, include in the trial function *adjustable parameters* $\alpha_1, \alpha_2, \ldots$ (i.e., $\mid \psi_0\rangle = |\psi_0(\alpha_1, \alpha_2, \ldots)\rangle$) which will account for the various possibilities of these unknown properties.

- Second, using (9.124), calculate the energy; this yields an expression which depends on the parameters $\alpha_1, \alpha_2, \ldots$:

$$E_0(\alpha_1, \alpha_2, \ldots) = \frac{\langle\psi_0(\alpha_1, \alpha_2, \ldots)|\hat{H}|\psi_0(\alpha_1, \alpha_2, \ldots)\rangle}{\langle\psi_0(\alpha_1, \alpha_2, \ldots)|\psi_0(\alpha_1, \alpha_2, \ldots)\rangle}. \tag{9.129}$$

  In most cases $|\psi_0(\alpha_1, \alpha_2, \ldots)\rangle$ will be assumed to be normalized; hence the denominator of this expression is equal to 1.

- Third, using (9.129), search for the minimum of $E_0(\alpha_1, \alpha_2, \ldots)$ by varying the adjustable parameters α_i until E_0 is minimized. That is, *minimize* $E(\alpha_1, \alpha_2, \ldots)$ with respect to α_1, $\alpha_2, \ldots$:

$$\frac{\partial E_0(\alpha_1, \alpha_2, \ldots)}{\partial \alpha_i} = \frac{\partial}{\partial \alpha_i} \frac{\langle\psi_0(\alpha_1, \alpha_2, \ldots)|\hat{H}|\psi_0(\alpha_1, \alpha_2, \ldots)\rangle}{\langle\psi_0(\alpha_1, \alpha_2, \ldots)|\psi_0(\alpha_1, \alpha_2, \ldots)\rangle} = 0, \tag{9.130}$$

  with $i = 1, 2, \ldots$. This gives the values of $(\alpha_{1_0}, \alpha_{2_0}, \ldots)$ that minimize E_0.

- Fourth, substitute these values of $(\alpha_{1_0}, \alpha_{2_0}, \ldots)$ into (9.129) to obtain the approximate value of the energy. The value $E_0(\alpha_{1_0}, \alpha_{2_0}, \ldots)$ thus obtained provides an upper bound for the exact ground state energy E_0. The exact ground state eigenstate $\mid \phi_0\rangle$ will then be approximated by the state $|\psi_0(\alpha_{1_0}, \alpha_{2_0}, \ldots)\rangle$.

What about the energies of the excited states? The variational method can also be used to find the approximate values for the energies of the first few excited states. For instance, to find the energy and eigenstate of the first excited state that will approximate E_1 and $\mid \phi_1\rangle$, we need to choose a trial function $\mid \psi_1\rangle$ that must be orthogonal to $\mid \psi_0\rangle$:

$$\langle\psi_1 \mid \psi_0\rangle = 0. \tag{9.131}$$

Then proceed as we did in the case of the ground state. That is, solve the variational equation (9.123) for $\mid \psi_1\rangle$:

$$\frac{\partial}{\partial \alpha_i} \frac{\langle\psi_1(\alpha_1, \alpha_2, \ldots)|\hat{H}|\psi_1(\alpha_1, \alpha_2, \ldots)\rangle}{\langle\psi_1(\alpha_1, \alpha_2, \ldots)|\psi_1(\alpha_1, \alpha_2, \ldots)\rangle} = 0 \qquad (i = 1, 2, \ldots). \tag{9.132}$$

Similarly, to evaluate the second excited state, we solve (9.123) for $|\psi_2\rangle$ and take into account the following two conditions:

$$\langle\psi_2 \mid \psi_0\rangle = 0, \qquad \langle\psi_2 \mid \psi_1\rangle = 0. \tag{9.133}$$

These conditions can be included in the variational problem by means of *Lagrange multipliers*, that is, by means of a constrained variational principle.

In this way, we can in principle evaluate any other excited state. However, the variational procedure becomes increasingly complicated as we deal with higher excited states. As a result, the method is mainly used to determine the ground state.

Remark

In those problems where the first derivative of the wave function is discontinuous at a given value of x, one has to be careful when using the expression

$$-\left\langle\psi\left|\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\right|\psi\right\rangle = -\frac{\hbar^2}{2m}\int_{-\infty}^{+\infty}\psi^*(x)\frac{d^2\psi(x)}{dx^2}\,dx. \tag{9.134}$$

A straightforward, careless use of this expression sometimes leads to a negative kinetic energy term (Problem 9.6 on page 593). One might instead consider using the following form:

$$\boxed{-\left\langle\psi\left|\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\right|\psi\right\rangle = \frac{\hbar^2}{2m}\int_{-\infty}^{+\infty}\left|\frac{d\psi(x)}{dx}\right|^2 dx.} \tag{9.135}$$

Note that (9.134) and (9.135) are identical; an integration by parts leads to

$$\int_{-\infty}^{+\infty}\left|\frac{d\psi(x)}{dx}\right|^2 dx = \psi^*(x)\frac{d\psi(x)}{dx}\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty}\psi^*(x)\frac{d^2\psi(x)}{dx^2}dx = -\int_{-\infty}^{+\infty}\psi^*(x)\frac{d^2\psi(x)}{dx^2}dx, \tag{9.136}$$

since $\psi^*(x)d\psi(x)/dx$ goes to zero as $x \longrightarrow \pm\infty$ (this is the case whenever $\psi(x)$ is a bound state, but not so when $\psi(x)$ is a plane wave).

What about the calculation of $\left\langle\psi\left|-(\hbar^2/(2m))\Delta\right|\psi\right\rangle$ in three dimensions? We might consider generalizing (9.135). For this, we need simply to invoke Gauss's theorem[3] to show that

$$\int\left(\vec{\nabla}\psi^*(\vec{r})\right)\cdot\left(\vec{\nabla}\psi(\vec{r})\right)d^3r = -\int\psi^*(\vec{r})\Delta\psi(\vec{r})\,d^3r. \tag{9.137}$$

To see this, an integration by parts leads to the following relation:

$$\int_S\psi^*(\vec{r})\vec{\nabla}\psi(\vec{r})\cdot d\vec{A} = \int_V\left[\left(\vec{\nabla}\psi^*(\vec{r})\right)\cdot\left(\vec{\nabla}\psi(\vec{r})\right) + \psi^*(\vec{r})\Delta\psi(\vec{r})\right]d^3r, \tag{9.138}$$

and since, as $S \longrightarrow \infty$, the surface integral $\int_S\psi^*(\vec{r})\vec{\nabla}\psi(\vec{r})\cdot d\vec{S}$ vanishes if $\psi(\vec{r})$ is a bound state, we recover (9.137). So the kinetic energy term (9.135) is given in three dimensions by

$$\boxed{-\left\langle\psi\left|\frac{\hbar^2}{2m}\Delta\right|\psi\right\rangle = \frac{\hbar^2}{2m}\int\left(\vec{\nabla}\psi^*(\vec{r})\right)\cdot\left(\vec{\nabla}\psi(\vec{r})\right)d^3r.} \tag{9.139}$$

---

[3]Gauss's theorem states that the surface integral of a vector $\vec{B}$ over a closed surface S is equal to the volume integral of the divergence of that vector integrated over the volume V enclosed by the surface S: $\int_S\vec{B}\cdot d\vec{S} = \int_V\vec{\nabla}\cdot\vec{B}\,dV$.

Example 9.4
Show that (9.123) is equivalent to the Schrödinger equation (9.122) .

Solution
Using (9.124), we can rewrite (9.123) as

$$\delta\left(\langle\psi \mid \hat{H} - E \mid \psi\rangle\right) = 0. \tag{9.140}$$

Since $\mid \psi\rangle$ is a complex function, we can view $\mid \psi\rangle$ and $\langle\psi \mid$ as two independent functions; hence we can carry out the variations over $|\delta\psi\rangle$ and $\langle\delta\psi|$ independently. Varying first over $\langle\delta\psi|$, equation (9.140) yields

$$\langle\delta\psi|\hat{H} - E \mid \psi\rangle = 0. \tag{9.141}$$

Since $\mid \psi\rangle$ is arbitrary, then (9.141) is equivalent to $\hat{H} \mid \psi\rangle = E \mid \psi\rangle$. The variation over $|\delta\psi\rangle$ leads to the same result. Namely, varying (9.140) over $|\delta\psi\rangle$, we get

$$\langle\psi|\hat{H} - E|\delta\psi\rangle = 0, \tag{9.142}$$

from which we obtain the complex conjugate equation $\langle\psi \mid \hat{H} = E\langle\psi \mid$, since $\hat{H}$ is Hermitian.

Example 9.5
Consider a one-dimensional harmonic oscillator. Use the variational method to estimate the energies of (a) the ground state, (b) the first excited state, and (c) the second excited state.

Solution
This simple problem enables us to illustrate the various aspects of the variational method within a predictable setting, because the exact solutions are known: $E_0 = \hbar\omega/2$, $E_1 = 3\hbar\omega/2$, $E_2 = 5\hbar\omega/2$.

(a) The trial function we choose for the ground state has to be even and smooth everywhere, it must vanish as $x \to \pm\infty$, and it must have no nodes. A Gaussian function satisfies these requirements. But what we are not sure about is its width. To account for this, we include in the trial function an adjustable scale parameter α:

$$\psi_0(x, \alpha) = Ae^{-\alpha x^2}; \tag{9.143}$$

A is a normalization constant. Using $\int_{-\infty}^{+\infty} x^{2n}e^{-ax^2}dx = \sqrt{\pi/a}\,1\cdot 3\cdot 5\cdots(2n-1)/(2a)^n$, we can show that A is given by $A = (2\alpha/\pi)^{1/4}$. The expression for $E_0(\alpha)$ is thus given by

$$\begin{aligned}\langle\psi_0|H|\psi_0\rangle &= A^2\int_{-\infty}^{+\infty} e^{-\alpha x^2}\left(\frac{-\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2x^2\right)e^{-\alpha x^2}dx \\ &= A^2\frac{\hbar^2\alpha}{m}\int_{-\infty}^{+\infty} e^{-2\alpha x^2}dx + A^2\left(\frac{1}{2}m\omega^2 - \frac{2\hbar^2\alpha^2}{m}\right)\int_{-\infty}^{+\infty} x^2e^{-2\alpha x^2}dx \\ &= \frac{\hbar^2\alpha}{m} + \frac{1}{4\alpha}\left(\frac{1}{2}m\omega^2 - \frac{2\hbar^2\alpha^2}{m}\right) = \frac{\hbar^2}{2m}\alpha + \frac{m\omega^2}{8\alpha}\end{aligned} \tag{9.144}$$

or

$$E_0(\alpha) = \frac{\hbar^2}{2m}\alpha + \frac{m\omega^2}{8\alpha}. \tag{9.145}$$

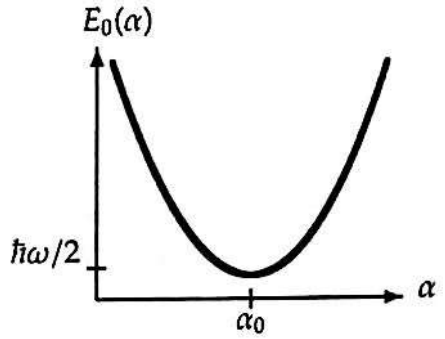


Figure 9.4: Shape of $E_0(\alpha) = \hbar^2\alpha/(2m) + m\omega^2/(8\alpha)$.

Its shape is displayed in Figure 9.4. The value of α_0, corresponding to the lowest point of the curve, can be obtained from the minimization of $E(\alpha)$ with respect to α,

$$\frac{\partial E_0(\alpha)}{\partial \alpha} = \frac{\hbar^2}{2m} - \frac{m\omega^2}{8\alpha^2} = 0, \tag{9.146}$$

yields $\alpha_0 = m\omega/(2\hbar)$ which, when inserted into (9.145) and (9.143), leads to

$$E_0(\alpha_0) = \frac{\hbar\omega}{2} \qquad \text{and} \qquad \psi_0(x,\alpha_0) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar}. \tag{9.147}$$

The ground state energy and wave function obtained by the variational method are identical to their exact counterparts.

(b) Let us now find the approximate energy E_1 for the first excited state. The trial function $\psi_1(x,\alpha)$ we need to select must be odd, it must vanish as $x \to \pm\infty$, it must have only one node, and it must be orthogonal to $\psi_0(x,\alpha_0)$ of (9.147). A candidate that satisfies these requirements is

$$\psi_1(x,\alpha) = Bxe^{-\alpha x^2}; \tag{9.148}$$

B is the normalization constant. We can show that $B = (32\alpha^3/\pi)^{1/4}$. Note that $\langle\psi_0 \mid \psi_1\rangle$ is zero,

$$\langle\psi_0 \mid \psi_1\rangle = B\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \int_{-\infty}^{+\infty} xe^{-\alpha x^2} e^{-m\omega x^2/2\hbar} dx = 0, \tag{9.149}$$

since the symmetric integration of an odd function is zero; $\psi_0(x)$ is even and $\psi_1(x)$ is odd.

Proceeding as we did for $E_0(\alpha)$, and since $\psi_1(x,\alpha)$ is normalized, we can show that

$$\begin{aligned} E_1(\alpha) &= \langle\psi_1(\alpha)|H|\psi_1(\alpha)\rangle = B^2 \int_{-\infty}^{+\infty} xe^{-\alpha x^2}\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2x^2\right] xe^{-\alpha x^2} dx \\ &= \frac{3\hbar^2}{2m}\alpha + \frac{3m\omega^2}{8\alpha}. \end{aligned} \tag{9.150}$$

The minimization of $E_1(\alpha)$ with respect to α (i.e., $\partial E_1(\alpha)/\partial\alpha = 0$) leads to $\alpha_0 = m\omega/2\hbar$. Hence the energy and the state of the first excited state are given by

$$E_1(\alpha_0) = \frac{3\hbar\omega}{2}, \qquad \psi_1(x,\alpha_0) = \left(\frac{4m^3\omega^3}{\pi\hbar^3}\right)^{1/4} xe^{-m\omega x^2/2\hbar}. \tag{9.151}$$

They are in full agreement with the exact expressions.

(c) The trial function

$$\psi_2(x,\alpha,\beta) = C(\beta x^2 - 1)e^{-\alpha x^2}, \tag{9.152}$$

which includes two adjustable parameters α and β, satisfies all the properties of the second excited state: even under parity, it vanishes as $x \to \pm\infty$ and has two nodes. The term $(\beta x^2 - 1)$ ensures that $\psi_2(x,\alpha,\beta)$ has two nodes $x = \pm 1/\sqrt{\beta}$ and the normalization constant C is given by

$$C = \left(\frac{2\alpha}{\pi}\right)^{1/4}\left[\frac{3\beta^2}{16\alpha^2} - \frac{\beta}{2\alpha} + 1\right]^{-1/2}. \tag{9.153}$$

The trial function $\psi_2(x,\alpha,\beta)$ must be orthogonal to both $\psi_0(x)$ and $\psi_1(x)$. First, notice that it is indeed orthogonal to $\psi_1(x)$, since $\psi_2(x,\alpha,\beta)$ is even while $\psi_1(x)$ is odd:

$$\langle\psi_1 \mid \psi_2\rangle = C\left(\frac{4m^3\omega^3}{\pi\hbar^3}\right)^{1/4}\int_{-\infty}^{+\infty} x(\beta x^2 - 1)e^{-\alpha x^2}e^{-m\omega x^2/2\hbar}dx = 0. \tag{9.154}$$

As for the orthogonality condition of $\psi_2(x)$ with $\psi_0(x)$, it can be written as

$$\begin{aligned}\langle\psi_0 \mid \psi_2\rangle &= \int_{-\infty}^{+\infty}\psi_0(x)\psi_2(x,\alpha,\beta)\,dx = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} C\int_{-\infty}^{+\infty}(\beta x^2 - 1)e^{-(m\omega/2\hbar+\alpha)x^2}dx \\ &= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} C\left[\frac{\beta}{2(m\omega/2\hbar+\alpha)} - 1\right]\sqrt{\frac{\pi}{m\omega/2\hbar+\alpha}} = 0.\end{aligned} \tag{9.155}$$

This leads to a useful condition between β and α:

$$\beta = \frac{m\omega}{\hbar} + 2\alpha. \tag{9.156}$$

Now let us focus on determining the energy $E_2(\alpha,\beta) = \langle\psi_2|\hat{H} \mid \psi_2\rangle$:

$$E_2(\alpha,\beta) = C^2\int_{-\infty}^{+\infty}(\beta x^2 - 1)e^{-\alpha x^2}\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2x^2\right](\beta x^2 - 1)e^{-\alpha x^2}dx. \tag{9.157}$$

After lengthy but straightforward calculations, we obtain

$$-\frac{\hbar^2}{2m}\left\langle\psi_2\left|\frac{d^2}{dx^2}\right|\psi_2\right\rangle = \frac{\hbar^2}{2m}\left(\alpha + \frac{\beta}{2} + \frac{7\beta^2}{16\alpha}\right)C^2\sqrt{\frac{\pi}{2\alpha}}, \tag{9.158}$$

$$\frac{1}{2}m\omega^2\langle\psi_2|x^2 \mid \psi_2\rangle = m\omega^2\left(\frac{15\beta^2}{128\alpha^3} - \frac{3\beta}{16\alpha^2} + \frac{1}{8\alpha}\right)C^2\sqrt{\frac{\pi}{2\alpha}}; \tag{9.159}$$

hence

$$E_2(\alpha,\beta) = C^2\sqrt{\frac{\pi}{2\alpha}}\left(\frac{\hbar^2\alpha}{2m} + \frac{\hbar^2\beta}{4m} + \frac{7\hbar^2\beta^2}{32m\alpha} + \frac{15m\beta^2\omega^2}{128\alpha^3} - \frac{3m\beta\omega^2}{16\alpha^2} + \frac{m\omega^2}{8\alpha}\right). \tag{9.160}$$

To extract the approximate value of E_2, we need to minimize $E_2(\alpha,\beta)$ with respect to α and to β: $\partial E_2(\alpha,\beta)/\partial\alpha = 0$ and $\partial E_2(\alpha,\beta)/\partial\beta = 0$. The two expressions we obtain will enable us to

extract (by solving a system of two linear equations with two unknowns) the values of α_0 and β_0 that minimize $E_2(\alpha, \beta)$. This method is lengthy and quite cumbersome; α_0 and β_0 have to satisfy the condition (9.156). We can, however, exploit this condition to come up with a much shorter approach: it consists of replacing the value of β as displayed in (9.156) into the energy relation (9.160), thereby yielding an expression that depends on a single parameter α:

$$E_2(\alpha) = \left(\frac{15\hbar^2\alpha}{18m} + \frac{9\hbar\omega}{8} + \frac{7m\omega^2}{16\alpha} + \frac{15m^3\omega^4}{128\hbar^2\alpha^3} + \frac{9m^2\omega^3}{32\hbar\alpha^2}\right)\left(\frac{3m^2\omega^2}{16\hbar^2\alpha^2} + \frac{m\omega}{4\hbar\alpha} + \frac{3}{4}\right)^{-1}; \tag{9.161}$$

in deriving this relation, we have substituted (9.156) into the expression for C as given by (9.153), which in turn is inserted into (9.160). In this way, we need to minimize E_2 with respect to one parameter only, α. This yields $\alpha_0 = m\omega/(2\hbar)$ which, when inserted into (9.156) leads to $\beta_0 = 2m\omega/\hbar$. Thus, the energy and wave function are given by

$$E_2(\alpha_0, \beta_0) = \frac{5}{2}\hbar\omega, \qquad \psi_2(x, \alpha_0, \beta_0) = \left(\frac{m\omega}{4\pi\hbar}\right)^{1/4}\left(\frac{2m\omega}{\hbar}x^2 - 1\right)e^{-\frac{m\omega}{2\hbar}x^2}. \tag{9.162}$$

These are identical with the exact expressions for the energy and the wave function.

Example 9.6
Use the variational method to estimate the ground state energy of the hydrogen atom.

Solution
The ground state wave function has no nodes and vanishes at infinity. Let us try

$$\psi(r, \theta, \phi) = e^{-r/a}, \tag{9.163}$$

where α is a scale parameter; there is no angular dependence of $\psi(r)$ since the ground state function is spherically symmetric. The energy is given by

$$E(\alpha) = \frac{\langle\psi \mid \hat{H} \mid \psi\rangle}{\langle\psi \mid \psi\rangle} = -\frac{\left\langle\psi\left|(\hbar^2/2m)\nabla^2 + e^2/r\right|\psi\right\rangle}{\langle\psi \mid \psi\rangle}, \tag{9.164}$$

where

$$\langle\psi \mid \psi\rangle = \int_0^{+\infty} r^2 e^{-2r/a} dr \int_0^{\pi} \sin\theta \, d\theta \int_0^{2\pi} d\phi = \pi\alpha^3 \tag{9.165}$$

and

$$-\left\langle\psi\left|\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{r}\right|\psi\right\rangle = -4\pi e^2 \int_0^{+\infty} r e^{-2r/a} dr = -\pi e^2\alpha^2. \tag{9.166}$$

To calculate the kinetic energy term, we may use (9.139)

$$-\left\langle\psi\left|\frac{\hbar^2}{2m}\nabla^2\right|\psi\right\rangle = \frac{\hbar^2}{2m}\int \left(\vec{\nabla}\psi^*(r)\right)\cdot\left(\vec{\nabla}\psi(r)\right) d^3r, \tag{9.167}$$

where

$$\vec{\nabla}\psi^*(r) = \vec{\nabla}\psi(r) = \frac{d\psi(r)}{dr}\hat{r} = -\frac{1}{\alpha}e^{-r/a}\hat{r}; \tag{9.168}$$

hence

$$-\left\langle \psi \left| \frac{\hbar^2}{2m}\nabla^2 \right| \psi \right\rangle = \frac{4\pi}{\alpha^2}\frac{\hbar^2}{2m}\int_0^{+\infty} r^2 e^{-2r/\alpha}\, dr = \frac{\hbar^2\pi}{2m}\alpha. \tag{9.169}$$

Inserting (9.165), (9.166), and (9.169) into (9.164), we obtain

$$E(\alpha) = \frac{\hbar^2}{2m\alpha^2} - \frac{e^2}{\alpha}. \tag{9.170}$$

Minimizing this relation with respect to α, $dE(\alpha)/d\alpha = -\hbar^2/(m\alpha_0^3) + e^2/\alpha_0^2 = 0$, we obtain $a_0 = \hbar^2/(me^2)$ which, when inserted into (9.170), leads to the ground state energy

$$E(\alpha_0) = -\frac{me^4}{2\hbar^2}. \tag{9.171}$$

This is the correct ground state energy for the hydrogen atom. The variational method has given back the correct energy because the trial function (9.163) happens to be identical with the exact ground state wave function. Note that the scale parameter $\alpha_0 = \hbar^2/(me^2)$ has the dimensions of length; it is equal to the Bohr radius.

9.4 The Wentzel–Kramers–Brillouin Method

The Wentzel–Kramers–Brillouin (WKB) method is useful for approximate treatments of systems with *slowly* varying potentials; that is, potentials which remain almost *constant* over a region of the order of the de Broglie wavelength. In the case of classical systems, this property is always satisfied since the wavelength of a classical system approaches zero. The WKB method can thus be viewed as a *semiclassical* approximation.

9.4.1 General Formalism

Consider the motion of a particle in a time-independent potential $V(\vec{r})$; the Schrödinger equation for the corresponding stationary state is

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r}) + V(\vec{r})\psi(\vec{r}) = E\psi(\vec{r}) \tag{9.172}$$

or

$$\nabla^2\psi(\vec{r}) + \frac{1}{\hbar^2}p^2(\vec{r})\psi(\vec{r}) = 0, \tag{9.173}$$

where $p(\vec{r})$ is the classical momentum at $\vec{r}$: $p(\vec{r}) = \sqrt{2m(E - V(\vec{r}))}$. If the particle is moving in a region where $V(\vec{r})$ is constant, the solution of (9.173) is of the form $\psi(\vec{r}) = Ae^{\pm i\vec{p}\cdot\vec{r}/\hbar}$. But how does one deal with those cases where $V(\vec{r})$ is not constant? The WKB method provides an approximate treatment for systems whose potentials, while not constant, are *slowly* varying functions of $\vec{r}$. That is, $V(\vec{r})$ is almost constant in a region which extends over several de Broglie wavelengths; we may recall that the de Broglie wavelength of a particle of mass m and energy E that is moving in a potential $V(\vec{r})$ is given by $\lambda = h/p = h/\sqrt{2m(E - V(\vec{r}))}$.

In essence, the WKB method consists of trying a solution to (9.173) in the following form:

$$\psi(\vec{r}) = A(\vec{r})e^{iS(\vec{r})/\hbar}, \tag{9.174}$$

where the amplitude $A(\vec{r})$ and the phase $S(\vec{r})$, which are real functions, are yet to be determined. Substituting (9.174) into (9.173) we obtain

$$A\left[\frac{\hbar^2}{A}\nabla^2 A - (\vec{\nabla}S)^2 + p^2(\vec{r})\right] + i\hbar\left[2(\vec{\nabla}A)\cdot(\vec{\nabla}S) + A\nabla^2 S\right] = 0. \tag{9.175}$$

The real and imaginary parts of this equation must vanish separately:

$$(\vec{\nabla}S)^2 = p^2(\vec{r}) = 2m(E - V(\vec{r})), \tag{9.176}$$

$$2(\vec{\nabla}A)\cdot(\vec{\nabla}S) + A\nabla^2 S = 0. \tag{9.177}$$

In deriving (9.176) we have neglected the term that contains $\hbar$ (i.e., $(\hbar^2/A)\nabla^2 A$), since it is small compared to $(\vec{\nabla}S)^2$ and to $p^2(\vec{r})$; $\hbar$ is considered to be very small for classical systems.

To illustrate the various aspects of the WKB method, let us consider the simple case of the *one-dimensional* motion of a single particle. We can thus reduce (9.176) and (9.177), respectively, to

$$\frac{dS}{dx} = \pm\sqrt{2m(E - V)} = \pm p(x), \tag{9.178}$$

$$2\left(\frac{d}{dx}\ln A\right)p(x) + \frac{d}{dx}p(x) = 0. \tag{9.179}$$

Let us find the solutions of (9.178) and (9.179). Integration of (9.178) yields

$$S(x) = \pm\int dx\,\sqrt{2m(E - V(x))} = \pm\int p(x)\,dx. \tag{9.180}$$

We can reduce (9.179) to

$$\frac{d}{dx}\left[2\ln A + \ln p(x)\right] = 0, \tag{9.181}$$

which in turn leads to

$$A(x) = \frac{C}{\sqrt{|p(x)|}}, \tag{9.182}$$

where C is an arbitrary constant. So (9.180) and (9.182) give, respectively, the phase $S(x)$ and amplitude $A(x)$ of the WKB wave function (9.174).

Inserting (9.182) and (9.180) into (9.174), we obtain two approximate solutions to equation (9.173):

$$\psi_\pm(x) = \frac{C_\pm}{\sqrt{|p(x)|}}\exp\left[\pm\frac{i}{\hbar}\int^x p(x')dx'\right]. \tag{9.183}$$

The amplitude of this wave function is proportional to $1/\sqrt{p(x)}$; hence the probability of finding the particle between x and $x + dx$ is proportional to $1/p(x)$. This is what we expect for a "classical" particle because the time it will take to travel a distance dx is proportional to the inverse of its speed (or its momentum).

We can now examine two separate cases corresponding to $E > V(x)$ and $E < V(x)$. First, let us consider the case $E > V(x)$, which is called the *classically allowed* region. Here $p(x)$ is a real function; the most general solution of (9.173) is a combination of $\psi_+(x)$ and $\psi_-(x)$:

$$\psi(x) = \frac{C_+}{\sqrt{p(x)}} \exp\left[\frac{i}{\hbar}\int^x p(x')dx'\right] + \frac{C_-}{\sqrt{p(x)}} \exp\left[-\frac{i}{\hbar}\int^x p(x')dx'\right]. \tag{9.184}$$

Second, in the case where $E < V(x)$, which is known as the *classically forbidden* region, the momentum $p(x)$ is imaginary and the exponents of (9.183) become real:

$$\psi(x) = \frac{C'_-}{\sqrt{|p(x)|}} \exp\left[-\frac{1}{\hbar}\int_x |p(x')|dx'\right] + \frac{C'_+}{\sqrt{|p(x)|}} \exp\left[\frac{1}{\hbar}\int^x |p(x')|dx'\right]. \tag{9.185}$$

Equations (9.184) and (9.185) give the system's wave function in the allowed and forbidden regions, respectively. But what about the structure of the wave function near the regions $E \simeq V(x)$? At the points x_i, we have $E = V(x_i)$; hence the momentum (9.178) vanishes, $p(x_i) = 0$. These points are called the *classical turning points*, because classically the particle stops at x_i and then turns back to resume its motion in the opposite direction. At these points the wave function (9.183) becomes infinite since $p(x_i) = 0$. One then needs to examine how to find the wave function at the turning points. Before looking into that, let us first study the condition of validity for the WKB approximation.

Validity of the WKB approximation

To obtain the condition of validity for the WKB method, let us examine the size of the various terms in (9.175), notably $A(\vec{\nabla}S)^2$ and $i\hbar A\nabla^2 S$. Since quantities of the order of $\hbar$ are too small in the classical limit, the quasi-classical region is expected to be given by the condition[4]

$$\left|\hbar\nabla^2 S\right| \ll (\vec{\nabla}S)^2, \tag{9.186}$$

which can be written in one dimension as

$$\hbar\left|\frac{S''}{S'^2}\right| \ll 1 \tag{9.187}$$

or

$$\left|\frac{d}{dx}\left(\frac{\hbar}{S'}\right)\right| \ll 1, \tag{9.188}$$

since $\nabla^2 S = d^2S/dx^2 = S''$ and $|\vec{\nabla}S| = dS(x)/dx = S'$. In what follows we are going to verify that this relation yields the condition of validity for the WKB approximation.

Since $S' = \pm p(x)$ (see (9.178)), we can reduce (9.188) to

$$\left|\frac{d\bar{\lambda}(x)}{dx}\right| \ll 1, \tag{9.189}$$

[4]The condition (9.186) can be found as follows. Substituting $\psi(\vec{r}) = e^{iS(\vec{r})/\hbar}$ into (9.173) and multiplying by $\hbar^2$, we get $i\hbar\nabla^2 S(\vec{r}) - (\vec{\nabla}S)^2 + p^2(\vec{r}) = 0$. In the classical limit, the term containing $\hbar$, $|i\hbar\nabla^2 S(\vec{r})|$, must be small compared to the terms that do not, $(\vec{\nabla}S)^2$; i.e., $|\hbar\nabla^2 S(\vec{r})| \ll (\vec{\nabla}S)^2$.

where $\bar{\lambda}(x) = \lambda(x)/(2\pi)$ and $\lambda(x)$ is the de Broglie wavelength of the particle:

$$\bar{\lambda}(x) = \frac{\hbar}{p(x)} = \frac{\hbar}{\sqrt{2m\,(E - V(x))}}. \tag{9.190}$$

The condition (9.189) means that the rate of change of the de Broglie wavelength is small (i.e., the wavelength of the particle must vary only slightly over distances of the order of its size). But this condition is always satisfied for classical systems. So the condition of validity for the WKB method is given by

$$\left|\frac{d\bar{\lambda}(x)}{dx}\right| = \left|\frac{d}{dx}\left(\frac{\hbar}{p(x)}\right)\right| \ll 1. \tag{9.191}$$

This condition clearly breaks down at the classical turning points, $E = V(x_i)$, since $p(x_i) = 0$; classically, the particle stops at $x = x_i$ and then moves in the opposite direction. As $p(x)$ becomes small, the wavelength (9.190) becomes large and hence violates the requirement that it remains small and varies only slightly; when $p(x)$ is too small, the condition (9.191) breaks down. So the WKB approximation is valid in both the allowed and forbidden regions but not at the classical turning points.

How does one specify the particle's wave function at $x = x_i$? Or how does one connect the allowed states (9.184) with their forbidden counterparts (9.185)? As we go through the classical turning point, from the allowed to the forbidden region and vice versa, we need to examine how to determine the particle's wave function everywhere and notably at the turning points. This is the most difficult issue of the WKB method, for it breaks down at the turning points. In the following section we are going to deal with this issue by solving the Schrödinger equation near and at $x = x_i$. We will do so by resorting to an approximation: we consider the potential to be given, near the turning points, by a straight line whose slope is equal to that of the potential at the turning point.

In what follows, we want to apply the WKB approximation to find the energy levels and the wave function of a particle moving in a potential well. We are going to show that the formulas giving the energy levels depend on whether or not the potential well has rigid walls. In fact, it even depends on the number of rigid walls the potential has. For this, we are going to consider three separate cases pertaining to the potential well with: no rigid walls, a single rigid wall, and two rigid walls.

9.4.2 Bound States for Potential Wells with No Rigid Walls

Consider a potential well that has no rigid walls as displayed in Figure 9.5. Here the classically forbidden regions are specified by $x < x_1$ and $x > x_2$, the classically allowed region by $x_1 < x < x_2$; x_1 and x_2 are the classical turning points. This is a suitable and simple example to illustrate the various aspects of the WKB method, notably how to determine the particle's wave function at the turning points. We will see how this method yields the Bohr–Sommerfeld quantization rule from which the bound state energies are to be extracted.

The WKB method applies everywhere in the three regions (1), (2), and (3), except near the two turning points $x = x_1$ and $x = x_2$ at which $E = V(x_1) = V(x_2)$. The WKB approximation to the wave function in regions (1) and (3) can be inferred from (9.185) and the approximation in region (2) from (9.184): the wave function must decay exponentially in regions (1) and (3)

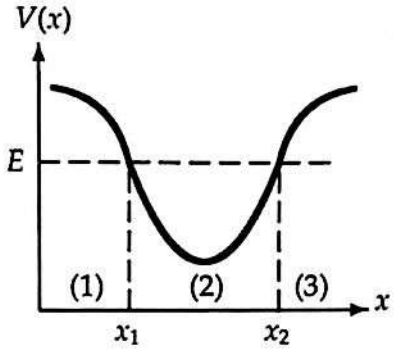


Figure 9.5: Potential with no rigid walls: regions (1) and (3) are classically forbidden, while (2) is classically allowed.

as $x \to -\infty$ and $x \to +\infty$, respectively, but must be oscillatory in region (2):

$$\psi_{1_{WKB}}(x) = \frac{C_1}{\sqrt{|p(x)|}} \exp\left[-\frac{1}{\hbar}\int_x^{x_1} |p(x')|\,dx'\right], \qquad x < x_1, \tag{9.192}$$

$$\psi_{2_{WKB}}(x) = \frac{C_2'}{\sqrt{p(x)}} \exp\left[\frac{i}{\hbar}\int_x p(x')dx'\right] + \frac{C_2''}{\sqrt{p(x)}} \exp\left[-\frac{i}{\hbar}\int_x p(x')dx'\right], \quad x_1 < x < x_2, \tag{9.193}$$

$$\psi_{3_{WKB}}(x) = \frac{C_3}{\sqrt{|p(x)|}} \exp\left[-\frac{1}{\hbar}\int_{x_2}^{x} |p(x')|\,dx'\right], \qquad x > x_2; \tag{9.194}$$

the coefficients C_1, C_2', C_2'', and C_3 have yet to be determined. For this, we must *connect* the solutions $\psi_1(x)$, $\psi_2(x)$, and $\psi_3(x)$ when passing from one region into another through the turning points $x = x_1$ and $x = x_2$ where the quasi-classical approximation ceases to be valid. That is, we need to connect $\psi_3(x)$ to $\psi_2(x)$ as we go from region (3) to (2), and then connect $\psi_1(x)$ to $\psi_2(x)$ as we go from (1) to (2). Since the WKB approximation breaks down at x_1 and x_2, we need to look for the exact solutions of the Schrödinger equation near x_1 and x_2.

9.4.2.1 Connection of $\psi_{3_{WKB}}(x)$ to $\psi_{2_{WKB}}(x)$

The WKB approximation to the wave function in region (2) can be inferred from (9.193):

$$\psi_{2_{WKB}}(x) = \frac{C_2'}{\sqrt{p(x)}} \exp\left[\frac{i}{\hbar}\int_x^{x_2} p(x')dx'\right] + \frac{C_2''}{\sqrt{p(x)}} \exp\left[-\frac{i}{\hbar}\int_x^{x_2} p(x')dx'\right], \quad x_1 < x < x_2; \tag{9.195}$$

this can be written as

$$\psi_{2_{WKB}}(x) = \frac{C_2}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_x^{x_2} p(x')dx' + \alpha\right), \qquad x_1 < x < x_2, \tag{9.196}$$

where α is a phase to be determined. Since the WKB approximation breaks down near the turning point x_2 (i.e., on both sides of $x = x_2$), we need to find a scheme for determining the wave function near x_2.

For this, let us now look for the exact solution of the Schrödinger equation near $x = x_2$. As mentioned above, if $|x - x_2|$ is small enough, within the region $|x - x_2|$, we can approximately represent the potential by a straight line whose slope is equal to that of the potential at the classical turning point $x = x_2$. That is, expanding $V(x)$ to first order around $x = x_2$, we obtain

$$V(x) \simeq V(x_2) + (x - x_2)\frac{dV(x)}{dx}\bigg|_{x=x_2} = E + (x - x_2)F_0, \tag{9.197}$$

where we have used the fact that $V(x_2) = E$ and where F_0 is given by $F_0 = \frac{dV(x)}{dx}\Big|_{x=x_2}$. Equation (9.197) means that $V(x)$ is approximated by a straight line $(x - x_2)F_0$, where F_0 is the slope of $V(x)$ at $x = x_2$. The Schrödinger equation for the potential (9.197) can be written as

$$\frac{d^2\psi(x)}{dx^2} - \frac{2mF_0}{\hbar^2}(x - x_2)\psi(x) = 0. \tag{9.198}$$

Using the change of variable

$$y = \left(\frac{2mF_0}{\hbar^2}\right)^{1/3}(x - x_2), \tag{9.199}$$

we can transform (9.198) into

$$\left(\frac{2mF_0}{\hbar^2}\right)^{2/3}\left[\frac{d^2\psi(y)}{dy^2} - y\psi(y)\right] = 0 \tag{9.200}$$

or

$$\frac{d^2\psi(y)}{dy^2} - y\psi(y) = 0. \tag{9.201}$$

This is a well-known differential equation whose solutions are usually expressed in terms of the Airy functions[5] $\mathrm{Ai}(y)$:

$$\psi(y) = A'\mathrm{Ai}(y) = \frac{A'}{\pi}\int_0^\infty \cos\left(\frac{z^3}{3} + yz\right)dz, \tag{9.202}$$

where A' is a normalization constant.

From the properties of the Airy function $\mathrm{Ai}(y) = 1/\pi \int_0^\infty \cos(z^3/3 + yz)dz$, the asymptotic behavior of $\mathrm{Ai}(y)$ is given for large positive and large negative values of y by

$$\mathrm{Ai}(y) \sim \begin{cases} \frac{1}{\sqrt{\pi}|y|^{1/4}}\sin\left[\frac{2}{3}(-y)^{3/2} + \frac{\pi}{4}\right], & y \ll 0, \\ \frac{1}{2\sqrt{\pi}y^{1/4}}\exp\left[-\frac{2}{3}y^{3/2}\right], & y \gg 0. \end{cases} \tag{9.203}$$

The asymptotic expression of (9.202) is therefore given for large positive and large negative values of y by

$$\psi(y) = \begin{cases} \frac{A'}{\sqrt{\pi}|y|^{1/4}}\sin\left[\frac{2}{3}(-y)^{3/2} + \frac{\pi}{4}\right], & y \ll 0, \\ \frac{A'}{2\sqrt{\pi}y^{1/4}}\exp\left[-\frac{2}{3}y^{3/2}\right], & y \gg 0. \end{cases} \tag{9.204}$$

[5]The solution to the differential equation $d^2\phi(y)/dy = y\phi(y)$ is given by the Airy function $\phi(y) = \mathrm{Ai}(y) = \frac{1}{\pi}\int_0^\infty \cos(z^3/3 + yz)dz$.

Since $F_0 > 0$ equation (9.199) implies that the cases $y \ll 0$ and $y \gg 0$ correspond to $x \ll x_2$ and $x \gg x_2$, respectively.

Now near the turning point $x = x_2$, (9.197) shows that $E - V(x) = -(x - x_2)F_0$; hence the square of the classical momentum $p^2(x)$ is given by

$$p^2(x) = 2m\,(E - V(x)) = -2m(x - x_2)F_0, \tag{9.205}$$

which is negative for $x > x_2$ and positive for $x < x_2$. Combining equations (9.199) and (9.205), we obtain

$$p^2(x) = -\,(2m\hbar F_0)^{2/3}\, y. \tag{9.206}$$

Now since $dx = (\hbar^2/(2mF_0))^{1/3}dy$ (see (9.199)), we use (9.206) to infer the following expression:

$$\frac{1}{\hbar}\int_x^{x_2} p(x')\,dx' = \frac{1}{\hbar}\,(2m\hbar F_0)^{1/3}\left(\frac{\hbar^2}{2mF_0}\right)^{1/3}\int_y^0 \sqrt{-y'}\,dy' = \int_y^0 \sqrt{-y'}\,dy' = \frac{2}{3}(-y)^{3/2}. \tag{9.207}$$

Inserting this into (9.204), we obtain

$$\psi(x) = \begin{cases} \frac{A}{\sqrt{p(x)}}\sin\left(\frac{1}{\hbar}\int_x^{x_2} p(x')dx' + \frac{\pi}{4}\right), & x \ll x_2, \\ \frac{A}{2\sqrt{|p(x)|}}\exp\left[-\frac{1}{\hbar}\int_x^{x_2} \left|p(x')\right| dx'\right], & x \gg x_2, \end{cases} \tag{9.208}$$

where $A = (2m\hbar F_0)^{1/6}A'/\sqrt{\pi}$. A comparison of (9.208a) with (9.196) and (9.208b) with (9.194) reveals that

$$A = 2C_3, \qquad C_2 = A, \qquad \alpha = \frac{\pi}{4}; \tag{9.209}$$

these expressions are known as the *connection formulas*, for they connect the WKB solutions at either side of a turning point. Since $\alpha = \pi/4$, $\psi_{2_{WKB}}(x)$ of (9.196) becomes

$$\psi_{2_{WKB}}(x) = \frac{C_2}{\sqrt{p(x)}}\sin\left(\frac{1}{\hbar}\int_x^{x_2} p(x')\,dx' + \frac{\pi}{4}\right). \tag{9.210}$$

9.4.2.2 Connection of $\psi_{1_{WKB}}(x)$ to $\psi_{2_{WKB}}(x)$

The WKB wave function for $x < x_1$ is given by (9.192); the WKB solution for $x > x_1$ can be inferred from (9.193):

$$\psi_{2_{WKB}}(x) = \frac{C_2'}{\sqrt{p(x)}}\exp\left[\frac{i}{\hbar}\int_{x_1}^{x} p(x')dx'\right] + \frac{C_2''}{\sqrt{p(x)}}\exp\left[-\frac{i}{\hbar}\int_x^{x_1} p(x')dx'\right], \quad x_1 < x < x_2, \tag{9.211}$$

which can be written as

$$\psi_{2_{WKB}}(x) = \frac{D}{\sqrt{p}}\sin\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx' + \beta\right). \tag{9.212}$$

Recall that near $x = x_1$ the WKB approximation breaks down.

The shape of the wave function near $x = x_1$ can, however, be found from an exact solution of the Schrödinger equation. For this, we proceed as we did for $x = x_2$. That is, we look for

the exact solution of the Schrödinger equation for small values of $|x - x_1|$. Expanding $V(x)$ near $x = x_1$, we obtain a Schrödinger equation similar to (9.201). Its solutions for $x < x_1$ and $x > x_1$ are given by expressions that are similar to (9.208b) and (9.208a) respectively:

$$\psi(x) = \begin{cases} \frac{E}{2\sqrt{|p(x)|}} \exp\left[\frac{1}{\hbar}\int_{x_1}^{x} |p(x')|\,dx'\right], & x \ll x_1, \\ \frac{E}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx' + \frac{\pi}{4}\right), & x \gg x_1. \end{cases} \tag{9.213}$$

Again, comparing (9.213a) with (9.192) and (9.213b) with (9.212), we obtain the other set of connection formulas:

$$E = 2C_1, \qquad E = D, \qquad \beta = \frac{\pi}{4}; \tag{9.214}$$

hence $\psi_2(x)$ of (9.212) becomes

$$\psi_{2_{WKB}}(x) = \frac{D}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx' + \frac{\pi}{4}\right). \tag{9.215}$$

9.4.2.3 Quantization of the Energy Levels of the Bound States

Since the two solutions (9.210) and (9.215) represent the same wave function in the same region, they must be equal:

$$\psi_{2_{WKB}}(x) = \frac{D}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx' + \frac{\pi}{4}\right) = \frac{C_2}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x}^{x_2} p(x')\,dx' + \frac{\pi}{4}\right). \tag{9.216}$$

This is an equation of the form $D\sin\theta_1 = C_2\sin\theta_2$. Its solutions must satisfy the following two relations. The first is $\theta_1 + \theta_2 = (n+1)\pi$, i.e.,

$$\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx' + \frac{\pi}{4}\right) + \left(\frac{1}{\hbar}\int_{x}^{x_2} p(x')\,dx' + \frac{\pi}{4}\right) = (n+1)\pi \tag{9.217}$$

or

$$\frac{1}{\hbar}\int_{x_1}^{x_2} p(x)\,dx = \left(n + \frac{1}{2}\right)\pi, \qquad n = 0, 1, 2, 3, \ldots; \tag{9.218}$$

and the second is

$$D = (-1)^n C_2. \tag{9.219}$$

Since the integral between the turning points $\int_{x_1}^{x_2} p(x)\,dx$ is equal to half the integral over a complete period of the quasi-classical motion of the particle, i.e., $\int_{x_1}^{x_2} p(x)\,dx = \frac{1}{2}\oint p(x)\,dx$, we can reduce (9.218) to

$$\oint p(x)\,dx = 2\int_{x_1}^{x_2} p(x)\,dx = \left(n + \frac{1}{2}\right)h, \qquad n = 0, 1, 2, 3, \ldots. \tag{9.220}$$

This relation determines the *quantized* (WKB) energy levels E_n of the bound states of a semi-classical system. It is similar to the *Bohr–Sommerfeld quantization rule*, which in turn is known

to represent an improved version of the Wilson–Sommerfeld rule $\oint p(x)\,dx = nh$, because the Wilson–Sommerfeld rule does not include the zero-point energy term $h/2$ (in the case of large values of n, where the classical approximation becomes reliable, we have $n + 1/2 \simeq n$; hence (9.220) reduces to $\oint p(x)\,dx = nh$). We can interpret this relation as follows: since the integral $\oint p(x)\,dx$ gives the area enclosed by the closed trajectory of the particle in the xp phase space, the condition (9.220) provides the mechanism for selecting, from the continuum of energy values of the semiclassical system, only those energies E_n for which the areas of the contours $p(x, E_n) = \sqrt{2m\,(E_n - V(x))}$ are equal to $(n + \frac{1}{2})h$:

$$\boxed{\oint p(x, E_n)\,dx = 2\int_{x_1}^{x_2} \sqrt{2m(E_n - V(x))}\,dx = \left(n + \frac{1}{2}\right)h,} \tag{9.221}$$

with $n = 0, 1, 2, 3, \ldots$. So in the xp phase space, the area between two successive bound states is equal to h: $\oint p(x, E_{n+1})\,dx - \oint p(x, E_n)\,dx = h$. Each single state therefore corresponds to an area h in the phase space. Note that the number n present in this relation is equal to the number of bound states; that is, the number of nodes of the wave function $\psi(x)$.

In summary, for a particle moving in a potential well like the one shown in Figure 9.5, the bound state energies can be extracted from the quantization rule (9.221) and the wave function is given in regions (1) and (3) by (9.192) and (9.194), respectively, and in region (2) either by (9.210) or (9.215). Combining the *connection relations* (9.209), (9.214), and (9.219) with the wave functions (9.192), (9.194), (9.210), and (9.215), we get the WKB approximation to the wave function:

$$\psi_{WKB}(x) = \begin{cases} \psi_{1_{WKB}}(x) = \dfrac{(-1)^n C_3}{\sqrt{|p(x)|}} \exp\left[-\frac{1}{\hbar}\int_x^{x_1} |p(x')|\,dx'\right], & x < x_1, \\ \psi_{3_{WKB}}(x) = \dfrac{C_3}{\sqrt{|p(x)|}} \exp\left[-\frac{1}{\hbar}\int_{x_2}^{x} |p(x')|\,dx'\right], & x > x_2. \end{cases} \tag{9.222}$$

In the region $x_1 < x < x_2$, $\psi_{2_{WKB}}(x)$ is given either by (9.210) or by (9.215)

$$\psi_{2_{WKB}}(x) = \begin{cases} \dfrac{2(-1)^n C_3}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')dx' + \frac{\pi}{4}\right), & x_1 < x < x_2, \\ \dfrac{2C_3}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x}^{x_2} p(x')dx' + \frac{\pi}{4}\right), & x_1 < x < x_2. \end{cases} \tag{9.223}$$

The coefficient C_3 has yet to be found from the normalization of $\psi_{WKB}(x)$. This is the wave function of the nth bound state.

Remark

An important application of the WKB method consists of using the quantization rule (9.221) to calculate the energy levels of central potentials. The energy of a particle of mass m bound in a central potential $V(r)$ is given by

$$E = \frac{p_r^2}{2m} + V_{\text{eff}}(r) = \frac{p_r^2}{2m} + V(r) + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}. \tag{9.224}$$

The particle is bound to move between the turning points r_1 and r_2 whose values are given by $E = V_{\text{eff}}(r_1) = V_{\text{eff}}(r_2)$ and its bound state energy levels can be obtained from

$$\int_{r_1}^{r_2} dr\, p_r(E, r) = \int_{r_1}^{r_2} dr \sqrt{2m\left(E - V(r) - \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right)} = \left(n + \frac{1}{2}\right)\pi\hbar, \tag{9.225}$$

where $n = 0, 1, 2, 3, \ldots$.

Example 9.7
Use the WKB method to estimate the energy levels of a one-dimensional harmonic oscillator.

Solution
The classical energy of a harmonic oscillator

$$E(x,p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 \tag{9.226}$$

leads to $p(E,x) = \pm\sqrt{2mE - m^2\omega^2x^2}$. At the turning points, x_{min} and x_{max}, the energy is given by $E = V(x) = \frac{1}{2}m\omega^2x^2$ where $x_{min} = -a$ and $x_{max} = a$ with $a = \sqrt{2E/(m\omega^2)}$. To obtain the quantized energy expression of the harmonic oscillator, we need to use the Bohr–Sommerfeld quantization rule (9.221):

$$\oint p\,dx = 2\int_{-a}^{a}\sqrt{2mE - m^2\omega^2x^2}dx = 4m\omega\int_0^a\sqrt{a^2 - x^2}dx. \tag{9.227}$$

Using the change of variable $x = a\sin\theta$, we have

$$\int_0^a\sqrt{a^2 - x^2}dx = a^2\int_0^{\pi/2}\cos^2\theta\,d\theta = \frac{a^2}{2}\int_0^{\pi/2}(1 + \cos 2\theta)d\theta = \frac{\pi a^2}{4} = \frac{\pi E}{2m\omega^2}; \tag{9.228}$$

hence

$$\oint p\,dx = \frac{2\pi E}{\omega}. \tag{9.229}$$

Since $\oint p\,dq = \left(n + \frac{1}{2}\right)h$ or $2\pi E/\omega = (n + 1/2)h$, we obtain

$$E_n^{WKB} = \left(n + \frac{1}{2}\right)\hbar\omega. \tag{9.230}$$

This expression is identical with the *exact* energy of the harmonic oscillator.

9.4.3 Bound States for Potential Wells with One Rigid Wall

Consider a particle moving in a potential well that has a rigid wall at $x = x_1$ (Figure 9.6); it is given by $V(x) = +\infty$ for $x < x_1$ and by a certain function $V(x)$ for $x > x_1$. The classically allowed region is specified by $x_1 < x < x_2$; x_1 and x_2 are the turning points.

To obtain the quantization rule which gives the bound state energy levels for this potential, we proceed as we did in obtaining (9.221). The WKB wave function in region $x_1 < x < x_2$ has an oscillatory form; it can be inferred from (9.212):

$$\psi_{WKB}(x) = \frac{A}{\sqrt{p(x)}}\sin\left(\frac{1}{\hbar}\int p(x')\,dx' + \alpha\right), \qquad x_1 \le x \le x_2, \tag{9.231}$$

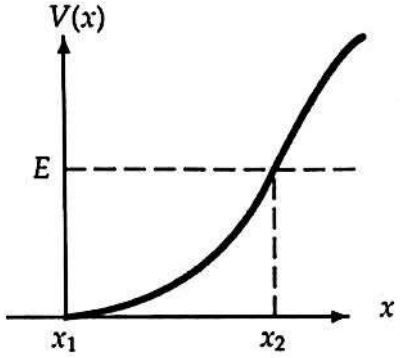


Figure 9.6: Potential well with one rigid wall located at $x = x_1$.

where α is a phase factor that needs to be specified. For this, we need to find the WKB wave function near the two turning points x_1 and x_2.

First, near x_2 (i.e., for $x \leq x_2$) we can determine the value of α as we did in obtaining (9.210). That is, expand $V(x)$ around $(x - x_2)$ and then match the WKB solutions at $x = x_2$; this leads to a phase factor $\alpha = \pi/4$ and hence

$$\psi_{WKB}(x) = \frac{B}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_x^{x_2} p(x')dx' + \frac{\pi}{4}\right), \qquad x_1 \leq x \leq x_2. \tag{9.232}$$

Second, since the wave function has to vanish at the rigid wall, $\psi_{WKB}(x_1) = 0$, the phase factor α must be zero; then (9.231) yields

$$\psi_{WKB}(x) = \frac{A}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx'\right), \qquad x_1 \leq x \leq x_2. \tag{9.233}$$

Now, since (9.232) and (9.233) represent the same wave function in the same region, the sum of their arguments must be equal to $(n + 1)\pi$ and $A = (-1)^n B$ (see Eq. (9.219)):

$$\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx'\right) + \left(\frac{1}{\hbar}\int_x^{x_2} p(x')\,dx' + \frac{\pi}{4}\right) = (n+1)\pi. \tag{9.234}$$

Thus, the quantization rule which gives the bound state energy levels for potential wells with one single rigid wall is given by

$$\boxed{\int_{x_1}^{x_2} p(x)\,dx = \left(n + \frac{3}{4}\right)\pi\hbar, \qquad n = 0, 1, 2, 3, lcdots.} \tag{9.235}$$

Remark

From the study carried out above, we may state that the phase factor α of the WKB solution (9.231) is in general equal to

- zero for turning points located at the rigid walls
- $\pi/4$ for turning points that are not located at the rigid walls.

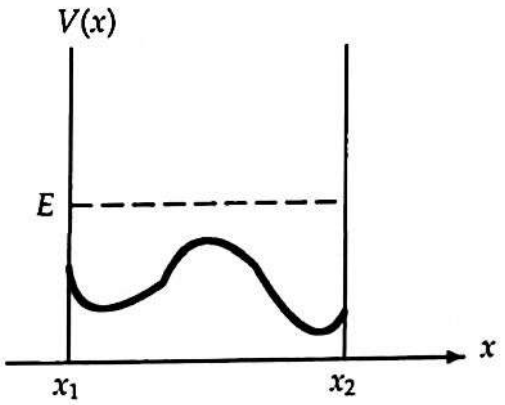


Figure 9.7: Potential well with two rigid walls located at x_1 and x_2.

9.4.4 Bound States for Potential Wells with Two Rigid Walls

Consider a potential well that has two rigid walls at $x = x_1$ and $x = x_2$. That is, as shown in Figure 9.7, $V(x)$ is infinite for $x \leq x_1$ and $x \geq x_2$ and given by a certain function $V(x)$ for $x_1 < x < x_2$. The wave function of a particle that is confined to move between the two rigid walls must vanish at the walls: $\psi(x_1) = \psi(x_2) = 0$.

To obtain the quantization rule which gives the bound state energy levels for this potential, we proceed as we did in obtaining (9.235). The WKB wave function has an oscillatory form in $x_1 < x < x_2$ and vanishes at both x_1 and x_2; the phase factor is zero at x_1 and x_2. By analogy with the procedure that led to (9.233), we can show that the WKB wave function in the vicinity of x_1 (i.e., in the region $x > x_1$) is given by

$$\psi_{WKB}(x) = \frac{A}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx'\right), \qquad x_1 < x < x_2, \tag{9.236}$$

and in the vicinity of x_2 (i.e., in the region $x < x_2$) it is given by

$$\psi_{WKB}(x) = \frac{B}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x}^{x_2} p(x')dx'\right), \qquad x_1 < x < x_2. \tag{9.237}$$

Note that the last two wave functions satisfy the correct boundary conditions at x_1 and x_2: $\psi_{WKB}(x_1) = \psi_{WKB}(x_2) = 0$.

Since equations (9.236) and (9.237) represent the same wave function in the same region, the sum of the arguments must then be equal to $(n+1)\pi$ and $A = (-1)^n B$ (see Eq. (9.219)):

$$\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx'\right) + \left(\frac{1}{\hbar}\int_{x}^{x_2} p(x')\,dx'\right) = (n+1)\pi; \tag{9.238}$$

hence the quantization rule for potential wells with two rigid walls is given by

$$\int_{x_1}^{x_2} p(x')\,dx' = (n+1)\,\pi\hbar, \qquad n = 0, 1, 2, 3, \ldots, \tag{9.239}$$

or by

$$\boxed{\int_{x_1}^{x_2} p(x)dx = n\pi\hbar, \qquad n = 1, 2, 3, \ldots.} \tag{9.240}$$

The only difference between (9.239) and (9.240) is in the minimum value of the quantum number n: the lowest value of n is $n = 0$ in (9.239) and $n = 1$ in (9.240).

Remark
In this section we have derived three quantization rules (9.221), (9.235), and (9.240); they provide the proper prescriptions for specifying the energy levels for potential wells with zero, one, and two rigid walls, respectively. These rules differ only in the numbers $\frac{1}{2}$, $\frac{3}{4}$, and 0 that are added to n. In the cases where n is large, which correspond to the semiclassical domain, these three quantization rules become identical; the semiclassical approximation is most accurate for large values of n.

Example 9.8
Use the WKB approximation to calculate the energy levels of a spinless particle of mass m moving in a one-dimensional box with walls at $x = 0$ and $x = L$.

Solution
This potential has two rigid walls, one at $x = 0$ and the other at $x = L$. To find the energy levels, we make use of the quantization rule (9.240). Since the momentum is constant within the well $p(E, x) = \sqrt{2mE}$, we can easily infer the WKB energy expression of the particle within the well. The integral is quite simple to calculate:

$$\int_0^L p\,dx = \sqrt{2mE}\int_0^L dx = L\sqrt{2mE}. \tag{9.241}$$

Now since $\int_0^L p\,dx = n\pi\hbar$ we obtain

$$L\sqrt{2mE_n^{WKB}} = n\pi\hbar; \tag{9.242}$$

hence

$$E_n^{WKB} = \frac{\pi^2\hbar^2}{2mL^2}n^2. \tag{9.243}$$

This is the exact value of the energy of a particle in an infinite well.

Example 9.9 (WKB method for the Coulomb potential)
Use the WKB approximation to calculate the energy levels of the s states of an electron that is bound to a Ze nucleus.

Solution
The electron moves in the Coulomb field of the Ze nucleus: $V(r) = -Ze^2/r$. Since the electron is bound to the nucleus, it can be viewed as moving between two rigid walls $0 \le r \le a$ with $E = V(a)$, $a = -Ze^2/E$; the energy of the electron is negative, $E < 0$.

The energy levels of the s states (i.e., $l = 0$) can thus be obtained from (9.240):

$$\int_0^a \sqrt{2m\left(E + \frac{Ze^2}{r}\right)}\,dr = n\pi\hbar. \tag{9.244}$$

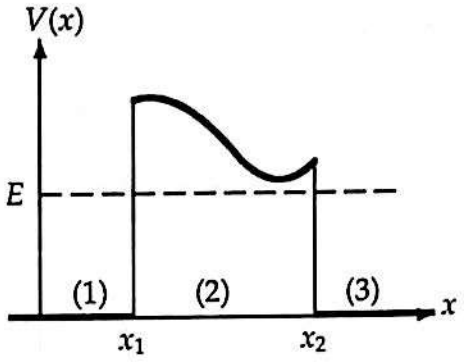


Figure 9.8: A potential barrier whose classically allowed regions are specified by $x < x_1$ and $x > x_2$ and the forbidden region by $x_1 < x < x_2$.

Using the change of variable $x = a/r$, we have

$$\int_0^a \sqrt{2m\left(E + \frac{Ze^2}{r}\right)}\,dr = \sqrt{-2mE}\int_0^a \sqrt{\frac{a}{r} - 1}\,dr = a\sqrt{-2mE}\int_0^1 \sqrt{\frac{1}{x} - 1}dx$$
$$= \frac{\pi}{2}a\sqrt{-2mE} = -\pi Ze^2\sqrt{-\frac{m}{2E}}. \tag{9.245}$$

In deriving this relation, we have used the integral $\int_0^1 \sqrt{1/x - 1}\,dx = \pi/2$; this can be easily obtained by the application of the residue theorem. Combining (9.244) and (9.245) we end up with

$$E_n = -\frac{mZ^2e^4}{2\hbar^2}\frac{1}{n^2} = -\frac{Z^2e^2}{2a_0}\frac{1}{n^2}, \tag{9.246}$$

where $a_0 = \hbar^2/(me^2)$ is the Bohr radius. This is the correct (Bohr) expression for the energy levels.

9.4.5 Tunneling through a Potential Barrier

Consider the motion of a particle of momentum $p_0 = \sqrt{2mE}$ incident from left onto a potential barrier $V(x)$, shown in Figure 9.8, with an energy E that is smaller than the potential's maximum value V_{max}.

Classically, the particle can in no way penetrate inside the barrier; hence it will get reflected backwards. Quantum mechanically, however, the probability corresponding to the particle's tunneling through the barrier and "emerging" to the right of the barrier is not zero. In what follows we want to use the WKB approximation to estimate the particle's probability of passing through the barrier.

In regions (1) and (3) of Figure 9.8 the particle is free:

$$\psi_1(x) = \psi_{incident}(x) + \psi_{reflected}(x) = Ae^{ip_0x/\hbar} + Be^{-ip_0x/\hbar}, \tag{9.247}$$
$$\psi_3(x) = \psi_{transmitted}(x) = Fe^{ip_0x/\hbar}, \tag{9.248}$$

where A, B, and F are the amplitudes of the incident, reflected, and transmitted waves, respectively; in region (3) we have outgoing waves only.

What about the wave function in the classically forbidden region (2)? The WKB method provides the answer. Since the particle energy is smaller than V_{max}, i.e., $E < V_{max}$, and if the potential $V(x)$ is a slowly varying function of x, the wave function in region (2) is given by the WKB approximation (see (9.185))

$$\psi_2(x) = \frac{C}{\sqrt{|p(x)|}} \exp\left[-\frac{1}{\hbar}\int_{x_1}^{x} |p(x')|dx'\right] + \frac{D}{\sqrt{|p(x)|}} \exp\left[\frac{1}{\hbar}\int_{x_1}^{x} |p(x')|dx'\right], \tag{9.249}$$

where $p(x) = i\sqrt{2m(V(x)-E)}$. The term $D/\sqrt{|p(x)|}\exp\left\{1/\hbar\int_{x_1}^{x}|p(x')|dx'\right\}$ increases exponentially when the barrier is very wide and is therefore unphysical. We shall be considering the case where the barrier is wide enough so that the approximation $D \simeq 0$ is valid; hence $\psi_2(x)$ becomes

$$\psi_2(x) = \frac{C}{\sqrt{|p(x)|}} \exp\left[-\frac{1}{\hbar}\int_{x_1}^{x} |p(x')|dx'\right]. \tag{9.250}$$

The probability corresponding to the particle's passage through the barrier is given by the transmission coefficient

$$T = \frac{v_{trans}}{v_{inc}} \frac{|\psi_{trans}(x)|^2}{|\psi_{inc}(x)|^2} = \frac{|F|^2}{|A|^2}, \tag{9.251}$$

since $v_{trans} = v_{inc}$ (the speeds of the incident and transmitted particles are equal). In what follows we are going to calculate the coefficient F in terms of A. For this, we need to use the continuity of the wave function and its derivative at x_1 and x_2. First, using (9.247) and (9.250), the continuity relations $\psi_1(x_1) = \psi_2(x_1)$ and $\psi_1'(x_1) = \psi_2'(x_1)$ lead, respectively, to

$$Ae^{ip_0x_1/\hbar} + Be^{-ip_0x_1/\hbar} = \frac{C}{\sqrt{a_1}}, \tag{9.252}$$

$$\frac{i}{\hbar}p_0(Ae^{ip_0x_1/\hbar} - Be^{-ip_0x_1/\hbar}) = -\frac{a_1}{\hbar\sqrt{a_1}}C, \tag{9.253}$$

where $a_1 = \sqrt{2m(V(x_1)-E)}$. The continuity of the wave function and its derivative at x_2, $\psi_2(x_2) = \psi_3(x_2)$, and $\psi_2'(x_2) = \psi_3'(x_2)$ lead to

$$\frac{C}{\sqrt{a_2}} \exp\left[-\frac{1}{\hbar}\int_{x_1}^{x_2} |p(x)|dx\right] = Fe^{ip_0x_2/\hbar} \tag{9.254}$$

$$-\frac{a_2}{\hbar\sqrt{a_2}}C \exp\left[-\frac{1}{\hbar}\int_{x_1}^{x_2} |p(x)|dx\right] = \frac{ip_0}{\hbar}Fe^{ip_0x_2/\hbar}, \tag{9.255}$$

where $a_2 = \sqrt{2m(V(x_2)-E)}$.

Adding (9.252) and (9.253) we get $C = 2A\sqrt{a_1}e^{ip_0x_1/\hbar}/(1 - a_1/ip_0)$ which, when inserted into (9.254), yields

$$\frac{F}{A} = \frac{2}{1 - a_1/ip_0}\sqrt{\frac{a_1}{a_2}}e^{ip_0(x_1-x_2)/\hbar} \exp\left[-\frac{1}{\hbar}\int_{x_1}^{x_2} |p(x)|dx\right], \tag{9.256}$$

which in turn leads to

$$\frac{|F|^2}{|A|^2} = \frac{4}{a_2/a_1 + a_1 a_2/p_0^2} \exp\left[-\frac{2}{\hbar}\int_{x_1}^{x_2} |p(x)|dx\right]. \tag{9.257}$$

The substitution of this expression into (9.251) finally yields an approximate value for the transmission coefficient through a potential barrier $V(x)$:

$$\boxed{T \sim e^{-2\gamma}, \qquad \gamma = \frac{1}{\hbar}\int_{x_1}^{x_2} \sqrt{2m(V(x) - E)}dx.} \tag{9.258}$$

Tunneling phenomena are common at the microscopic scale; they occur within nuclei, within atoms, and within solids. In nuclear physics, for instance, there are nuclei that decay into an α-particle (helium nucleus with $Z = 2$) and a daughter nucleus. This process can be viewed as the tunneling of an α-particle through the potential (Coulomb) barrier between the α-particle and the daughter nucleus; once formed inside the nucleus, the α-particle cannot escape unless it tunnels through (penetrates) the Coulomb barrier surrounding it. Tunneling also occurs within metals; when a metal is subject to an external electric field, electrons can be emitted from the metal. This is known as *cold emission*; we will study it in Example 9.10.

Example 9.10
Use the WKB approximation to estimate the transmission coefficient of a particle of mass m and energy E moving in the following potential barrier:

$$V(x) = \begin{cases} 0, & x < 0, \\ V_0 - \lambda x & x > 0. \end{cases}$$

Solution
The transmission coefficient is given by (9.258), where $x_1 = 0$ and the value of x_2, which can be obtained from the relation $V_0 - \lambda x_2 = E$, is given by $x_2 = (V_0 - E)/\lambda$. Setting the values of x_1 and x_2 into (9.258), and since $V(x) - E = (V_0 - E) - \lambda x$, we get

$$\begin{aligned} \gamma &= \frac{1}{\hbar}\int_{x_1}^{x_2} \sqrt{2m(V(x) - E)}\,dx = \frac{\sqrt{2m}}{\hbar}\int_0^{(V_0-E)/\lambda} \sqrt{V_0 - E - \lambda x}\,dx \\ &= \frac{2\sqrt{2m}}{3\hbar\lambda}(V_0 - E)^{3/2}. \end{aligned} \tag{9.259}$$

The transmission coefficient is thus given by

$$T \sim e^{-2\gamma} = \exp\left\{-\frac{4\sqrt{2m}}{3\hbar\lambda}(V_0 - E)^{3/2}\right\}. \tag{9.260}$$

This problem is useful for the study of *cold emission* of electrons from metals. In the absence of any external electric field, the electrons are bound by a potential of the type $V(x) = V_0$ for $x > 0$, known as the work function of the metal. When we turn on an external electric field $\mathcal{E}$, the potential seen by the electron is no longer V_0 but $V(x) = V_0 - e\mathcal{E}x$. This potential barrier

has a width through which the electrons can escape: every electron of energy $E \geq e\mathcal{E}x$ can escape. The quantity $e\mathcal{E}x_2$, where $x_2 = (V_0 - E)/\lambda$, is known as the work function of the metal; the width of the potential barrier of the metal is given by $0 < x < x_2$.

9.5 Concluding Remarks

In this chapter we have studied three approximation methods that apply to stationary Hamiltonians. As we saw, approximation methods offer efficient, short ways for obtaining energy levels that are, at times, identical with the exact results. For instance, in the calculation of the energy levels of the harmonic oscillator and the hydrogen atom, we have seen in a number of solved examples how the variational method and the WKB method lead to the correct energies *without* resorting to solve the Schrödinger equation; the approximation methods deal merely with the solution of a few simple integrals. In Chapters 4 and 7, however, we have seen that, to solve the Schrödinger equation for the harmonic oscillator and for the hydrogen atom, one has to carry out lengthy, laborious calculations.

Approximation methods offer, in general, powerful economical prescriptions for determining reliable results for systems that cannot be solved exactly. In the next chapter we are going to study approximation methods that apply to time-dependent processes such as atomic transitions, decays, and so on.

9.6 Solved Problems

The topic of approximation methods touches on almost all areas of quantum mechanics, ranging from one- to three-dimensional problems, as well as on the various aspects of the formalism of quantum mechanics.

Problem 9.1
Using first-order perturbation theory, calculate the energy of the nth excited state for a spinless particle of mass m moving in an infinite potential well of length $2L$, with walls at $x = 0$ and $x = 2L$:

$$V(x) = \begin{cases} 0, & 0 \leq x \leq 2L, \\ \infty, & \text{otherwise}, \end{cases}$$

which is modified at the bottom by the following two perturbations:
(a) $V_p(x) = \lambda V_0 \sin(\pi x/2L)$; (b) $V_p(x) = \lambda V_0 \delta(x - L)$, where $\lambda \ll 1$.

Solution
The exact expressions of the energy levels and of the wave functions for this potential are given by

$$E_n = \frac{\hbar^2 \pi^2}{8mL^2} n^2, \qquad \psi_n(x) = \frac{1}{\sqrt{L}} \sin\left(\frac{n\pi x}{2L}\right). \tag{9.261}$$

According to perturbation theory, the energy of the nth state is given to first order by

$$E_n = \frac{\hbar^2\pi^2}{8mL^2}n^2 + E_n^{(1)}, \tag{9.262}$$

where

$$E_n^{(1)} = \langle\psi_n|V_p(x)\,|\,\psi_n\rangle = \frac{1}{L}\int_0^{2L}\sin^2\left(\frac{n\pi x}{2L}\right)V_p(x)\,dx. \tag{9.263}$$

(a) Using the relation

$$\int \cos nx \sin mx dx = -\frac{\cos(m-n)x}{2(m-n)} - \frac{\cos(m+n)x}{2(m+n)}, \qquad m \neq \pm n, \tag{9.264}$$

along with (9.263), we can calculate $E_n^{(1)}$ for $V_p(x) = \lambda V_0 \sin(\pi x/2L)$ as follows:

$$\begin{aligned} E_n^{(1)} &= \frac{\lambda V_0}{L}\int_0^{2L}\sin^2\left(\frac{n\pi x}{2L}\right)\sin\left(\frac{\pi x}{2L}\right)dx \\ &= \frac{\lambda V_0}{2L}\int_0^{2L}\left[1-\cos\left(\frac{n\pi x}{L}\right)\right]\sin\left(\frac{\pi x}{2L}\right)dx \\ &= \frac{\lambda V_0}{\pi}\left\{-\cos\left(\frac{\pi x}{2L}\right) + \frac{\cos[(1-2n)\pi x/(2L)]}{2(1-2n)} + \frac{\cos[(1+2n)\pi x/(2L)]}{2(1+2n)}\right\}\Bigg|_0^{2L} \\ &= \frac{2\lambda V_0}{\pi}\frac{4n^2}{4n^2-1}. \end{aligned} \tag{9.265}$$

Thus, the energy (9.262) would become

$$E_n = \frac{\hbar^2\pi^2}{8mL^2}n^2 + \frac{2\lambda V_0}{\pi}\frac{4n^2}{4n^2-1}. \tag{9.266}$$

(b) In the case of $V_p(x) = \lambda V_0\delta(x-L)$, (9.263) leads to

$$E_n^{(1)} = \frac{\lambda V_0}{L}\int_0^{2L}\sin^2\left(\frac{n\pi x}{2L}\right)\delta(x-L)dx = \frac{\lambda V_0}{L}\sin^2\left(\frac{n\pi}{2}\right); \tag{9.267}$$

hence, depending on whether the quantum number n is even or odd, we have

$$E_n = \frac{\hbar^2\pi^2}{8mL^2}n^2 + \begin{cases} 0 & \text{if } n \text{ is even,} \\ \lambda V_0/L, & \text{if } n \text{ is odd.} \end{cases} \tag{9.268}$$

Problem 9.2

Consider a system whose Hamiltonian is given by $\hat{H} = E_0\begin{pmatrix} 1+\lambda & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 3 & -2\lambda \\ 0 & 0 & -2\lambda & 7 \end{pmatrix}$, where $\lambda \ll 1$.

(a) By decomposing this Hamiltonian into $\hat{H} = \hat{H}_0 + \hat{H}_p$, find the eigenvalues and eigenstates of the unperturbed Hamiltonian $\hat{H}_0$.

(b) Diagonalize $\hat{H}$ to find the exact eigenvalues of $\hat{H}$; expand each eigenvalue to the second power of λ.

(c) Using first- and second-order nondegenerate perturbation theory, find the approximate eigenergies of $\hat{H}$ and the eigenstates to first order. Compare these with the exact values obtained in (b).

Solution

(a) The matrix of $\hat{H}$ can be separated as follows:

$$H = \hat{H}_0 + \hat{H}_p = E_0 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix} + E_0 \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\lambda \\ 0 & 0 & -2\lambda & 0 \end{pmatrix}. \tag{9.269}$$

Notice that $\hat{H}_0$ is already diagonal; hence its eigenvalues are given by

$$E_1^{(0)} = E_0, \quad E_2^{(0)} = 8E_0, \quad E_3^{(0)} = 3E_0, \quad E_4^{(0)} = 7E_0, \tag{9.270}$$

and its eigenstates by

$$|\phi_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |\phi_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\phi_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |\phi_4\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{9.271}$$

(b) The diagonalization of $\hat{H}$ leads to the following secular equation:

$$\begin{vmatrix} (1+\lambda)E_0 - E & 0 & 0 & 0 \\ 0 & 8E_0 - E & 0 & 0 \\ 0 & 0 & 3E_0 - E & -2\lambda E_0 \\ 0 & 0 & -2\lambda E_0 & 7E_0 - E \end{vmatrix} = 0 \tag{9.272}$$

or

$$(E_0 + \lambda E_0 - E)(8E_0 - E)\left[(3E_0 - E)(7E_0 - E) - 4\lambda^2 E_0^2\right] = 0, \tag{9.273}$$

which in turn leads to the following *exact* eigenenergies:

$$E_1 = (1+\lambda)E_0, \quad E_2 = 8E_0, \quad E_3 = (5 - 2\sqrt{1+\lambda^2})E_0, \quad E_4 = (5 + 2\sqrt{1+\lambda^2})E_0. \tag{9.274}$$

Since $\lambda \ll 1$ we can expand $\sqrt{1+\lambda^2}$ to second order in λ: $\sqrt{1+\lambda^2} \simeq 1 + \lambda^2/2$. Hence E_3 and E_4 are given to second order in λ by

$$E_3 \simeq (3 - \lambda^2)E_0, \qquad E_4 \simeq (7 + \lambda^2)E_0. \tag{9.275}$$

(c) From nondegenerate perturbation theory, we can write the first-order corrections to the energies as follows:

$$E_1^{(1)} = \langle \phi_1 | \hat{H}_p | \phi_1 \rangle = E_0 (1\,0\,0\,0) \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\lambda \\ 0 & 0 & -2\lambda & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \lambda E_0. \tag{9.276}$$

Similarly, we can verify that the second, third, and fourth eigenvalues have no first-order corrections:

$$E_2^{(1)} = \langle \phi_2 \mid \hat{H}_p \mid \phi_2 \rangle = 0, \quad E_3^{(1)} = \langle \phi_3 \mid \hat{H}_p \mid \phi_3 \rangle = 0, \quad E_4^{(1)} = \langle \phi_4 \mid \hat{H}_p \mid \phi_4 \rangle = 0. \tag{9.277}$$

Let us now consider the second-order corrections to the energy. From nondegenerate perturbation theory, we have

$$E_1^{(2)} = \sum_{m=2,3,4} \frac{\left|\langle \phi_m | \hat{H}_p \mid \phi_1 \rangle\right|^2}{E_1^{(0)} - E_m^{(0)}} = 0, \tag{9.278}$$

since $\langle \phi_2 \mid \hat{H}_p \mid \phi_1 \rangle = \langle \phi_3 \mid \hat{H}_p \mid \phi_1 \rangle = \langle \phi_4 \mid \hat{H}_p \mid \phi_1 \rangle = 0$. Similarly, we can verify that

$$E_2^{(2)} = \sum_{m=1,3,4} \frac{\left|\langle \phi_m | \hat{H}_p \mid \phi_2 \rangle\right|^2}{E_2^{(0)} - E_m^{(0)}} = 0 \tag{9.279}$$

and

$$E_3^{(2)} = \sum_{m=1,2,4} \frac{\left|\langle \phi_m | \hat{H}_p \mid \phi_3 \rangle\right|^2}{E_3^{(0)} - E_m^{(0)}} = \frac{\left|\langle \phi_4 \mid \hat{H}_p \mid \phi_3 \rangle\right|^2}{E_3^{(0)} - E_4^{(0)}} = \frac{(-2\lambda E_0)^2}{(3-7)E_0} = -\lambda^2 E_0, \tag{9.280}$$

because

$$\langle \phi_4 \mid \hat{H}_p \mid \phi_3 \rangle = E_0 (0\,0\,0\,1) \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\lambda \\ 0 & 0 & -2\lambda & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = -2\lambda E_0. \tag{9.281}$$

Similarly, since

$$\langle \phi_3 \mid \hat{H}_p \mid \phi_4 \rangle = E_0 (0\,0\,1\,0) \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\lambda \\ 0 & 0 & -2\lambda & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = -2\lambda E_0, \tag{9.282}$$

we can ascertain that

$$E_4^{(2)} = \sum_{m=1,2,3} \frac{\left|\langle \phi_m | \hat{H}_p \mid \phi_4 \rangle\right|^2}{E_4^{(0)} - E_m^{(0)}} = \frac{\left|\langle \phi_3 \mid \hat{H}_p \mid \phi_4 \rangle\right|^2}{E_4^{(0)} - E_3^{(0)}} = \frac{(-2\lambda E_0)^2}{(7-3)E_0} = \lambda^2 E_0. \tag{9.283}$$

Now, combining (9.276)-(9.283), we infer that the values of the energies to second-order nondegenerate perturbation theory are given by

$$E_1 = E_1^{(0)} + E_1^{(1)} + E_1^{(2)} = (1+\lambda)E_0, \tag{9.284}$$

$$E_2 = E_2^{(0)} + E_2^{(1)} + E_2^{(2)} = 8E_0, \tag{9.285}$$

$$E_3 = E_3^{(0)} + E_3^{(1)} + E_3^{(2)} = (3-\lambda^2)E_0, \tag{9.286}$$

$$E_4 = E_4^{(0)} + E_4^{(1)} + E_4^{(2)} = (7+\lambda^2)E_0. \tag{9.287}$$

All these values are identical with their corresponding exact expressions (9.274) and (9.275).

Finally, according to (9.19), the first-order corrections to the eigenstates are given by

$$|\psi_n^{(1)}\rangle = \sum_{m\neq n} \frac{\langle\phi_m|\hat{H}_p|\phi_n\rangle}{E_n^{(0)} - E_m^{(0)}} |\phi_m\rangle, \tag{9.288}$$

and hence

$$|\psi_1^{(1)}\rangle = \sum_{m=2,3,4} \frac{\langle\phi_m|\hat{H}_p|\phi_1\rangle}{E_1^{(0)} - E_m^{(0)}} |\phi_m\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{9.289}$$

Similarly, we can show that $|\psi_2^{(1)}\rangle$ is also given by a zero column matrix, but $|\psi_3^{(1)}\rangle$ and $|\psi_4^{(1)}\rangle$ are not:

$$|\psi_3^{(1)}\rangle = \sum_{m=1,2,4} \frac{\langle\phi_m|\hat{H}_p|\phi_3\rangle}{E_3^{(0)} - E_m^{(0)}} |\phi_m\rangle = \frac{\langle\phi_4|\hat{H}_p|\phi_3\rangle}{E_3^{(0)} - E_4^{(0)}} |\phi_4\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \lambda/2 \end{pmatrix}, \tag{9.290}$$

$$|\psi_4^{(1)}\rangle = \sum_{m=1,2,3} \frac{\langle\phi_m|\hat{H}_p|\phi_4\rangle}{E_4^{(0)} - E_m^{(0)}} |\phi_m\rangle = \frac{\langle\phi_3|\hat{H}_p|\phi_4\rangle}{E_4^{(0)} - E_3^{(0)}} |\phi_3\rangle = \begin{pmatrix} 0 \\ 0 \\ -\lambda/2 \\ 0 \end{pmatrix}. \tag{9.291}$$

Finally, the states are given to first order by $|\psi_n\rangle = |\phi_n\rangle + |\psi_n^{(1)}\rangle$:

$$|\psi_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |\psi_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\psi_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \lambda/2 \end{pmatrix}, \quad |\psi_4\rangle = \begin{pmatrix} 0 \\ 0 \\ -\lambda/2 \\ 1 \end{pmatrix}. \tag{9.292}$$

Problem 9.3

(a) Find the exact energies and wave functions of the ground and first excited states and specify their degeneracies for the infinite cubic potential well

$$V(x,y,z) = \begin{cases} 0 & \text{if } 0 < x < L, 0 < y < L, 0 < z < L, \\ \infty & \text{otherwise.} \end{cases}$$

Now add the following perturbation to the infinite cubic well:

$$\hat{H}_p = V_0 L^3 \delta\left(x - \frac{L}{4}\right)\delta\left(y - \frac{3L}{4}\right)\delta\left(z - \frac{L}{4}\right).$$

(b) Using first-order perturbation theory, calculate the energy of the ground state.

(c) Using first-order (degenerate) perturbation theory, calculate the energy of the first excited state.

Solution

The energy and wave function for an infinite, cubic potential well of size L are given by

$$E^{exact}_{n_x,n_y,n_z} = \frac{\pi^2\hbar^2}{2mL^2}\left(n_x^2+n_y^2+n_z^2\right), \tag{9.293}$$

$$\phi_{n_x,n_y,n_z}(x,y,z) = \sqrt{\frac{8}{L^3}}\sin\left(\frac{\pi n_x}{L}x\right)\sin\left(\frac{\pi n_y}{L}y\right)\sin\left(\frac{\pi n_z}{L}z\right). \tag{9.294}$$

(a) The ground state is not degenerate; its exact energy and wave function are

$$E^{exact}_{111} = \frac{3\pi^2\hbar^2}{2mL^2}, \qquad \phi_{111}(x,y,z) = \sqrt{\frac{8}{L^3}}\sin\left(\frac{\pi}{L}x\right)\sin\left(\frac{\pi}{L}y\right)\sin\left(\frac{\pi}{L}z\right). \tag{9.295}$$

The first excited state is threefold degenerate: $\phi_{112}(x,y,z)$, $\phi_{121}(x,y,z)$, and $\phi_{211}(x,y,z)$ correspond to the same energy, $E^{exact}_{112} = E^{exact}_{121} = E^{exact}_{211} = 3\pi^2\hbar^2/(mL^2)$.

(b) The first-order correction to the ground state energy is given by

$$\begin{aligned} E_1^{(1)} &= \langle\phi_{111}\,|\,\hat{H}_p\,|\,\phi_{111}\rangle \\ &= 8V_0\int_0^L \delta\left(x-\frac{L}{4}\right)\sin^2\left(\frac{\pi}{L}x\right)dx\int_0^L \delta\left(y-\frac{3L}{4}\right)\sin^2\left(\frac{\pi}{L}y\right)dy \\ &\quad\times\int_0^L \delta\left(z-\frac{L}{4}\right)\sin^2\left(\frac{\pi}{L}z\right)dz = 8V_0\sin^2\left(\frac{\pi}{4}\right)\sin^2\left(\frac{3\pi}{4}\right)\sin^2\left(\frac{\pi}{4}\right) \\ &= V_0. \end{aligned} \tag{9.296}$$

Thus, the ground state energy is given to first-order perturbation by

$$E_0 = \frac{3\pi^2\hbar^2}{2mL^2} + V_0. \tag{9.297}$$

(c) To calculate the energy of the first excited state to first order, we need to use degenerate perturbation theory. The values of this energy are equal to $3\pi^2\hbar^2/(mL^2)$ plus the eigenvalues of the matrix

$$\begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{pmatrix}, \tag{9.298}$$

with $V_{nm} = \langle n\,|\,\hat{H}_p\,|\,m\rangle$, and where the following notations are used:

$$|\,1\rangle = \phi_{211}(x,y,z) = \sqrt{\frac{8}{L^3}}\sin\left(\frac{2\pi}{L}x\right)\sin\left(\frac{\pi}{L}y\right)\sin\left(\frac{\pi}{L}z\right), \tag{9.299}$$

$$|\,2\rangle = \phi_{121}(x,y,z) = \sqrt{\frac{8}{L^3}}\sin\left(\frac{\pi}{L}x\right)\sin\left(\frac{2\pi}{L}y\right)\sin\left(\frac{\pi}{L}z\right), \tag{9.300}$$

$$|\,3\rangle = \phi_{112}(x,y,z) = \sqrt{\frac{8}{L^3}}\sin\left(\frac{\pi}{L}x\right)\sin\left(\frac{\pi}{L}y\right)\sin\left(\frac{2\pi}{L}z\right). \tag{9.301}$$

The calculations of the terms V_{nm} are lengthy but straightforward. Let us show how to calculate two such terms. First, V_{11} can be calculated in analogy to (9.296):

$$\begin{aligned} V_{11} &= 8V_0 \int_0^L \delta\left(x-\frac{L}{4}\right)\sin^2\left(\frac{2\pi}{L}x\right)dx \int_0^L \delta\left(y-\frac{3L}{4}\right)\sin^2\left(\frac{\pi}{L}y\right)dy \\ &\quad \times \int_0^L \delta\left(z-\frac{L}{4}\right)\sin^2\left(\frac{\pi}{L}z\right)dz = 8V_0 \sin^2\left(\frac{\pi}{2}\right)\sin^2\left(\frac{3\pi}{4}\right)\sin^2\left(\frac{\pi}{4}\right) \\ &= 2V_0; \end{aligned} \tag{9.302}$$

V_{12} and V_{13} are given by

$$\begin{aligned} V_{12} &= 8V_0 \int_0^L \delta\left(x-\frac{L}{4}\right)\sin\left(\frac{2\pi}{L}x\right)\sin\left(\frac{\pi}{L}x\right)dx \int_0^L \delta\left(y-\frac{3L}{4}\right)\sin\left(\frac{\pi}{L}y\right) \\ &\quad \times \sin\left(\frac{2\pi}{L}y\right)dy \int_0^L \delta\left(z-\frac{L}{4}\right)\sin^2\left(\frac{\pi}{L}z\right)dz = -2V_0, \end{aligned} \tag{9.303}$$

$$\begin{aligned} V_{13} &= 8V_0 \int_0^L \delta\left(x-\frac{L}{4}\right)\sin\left(\frac{2\pi}{L}x\right)\sin\left(\frac{\pi}{L}x\right)dx \int_0^L \delta\left(y-\frac{3L}{4}\right)\sin^2\left(\frac{\pi}{L}y\right)dy \\ &\quad \times \int_0^L \delta\left(z-\frac{L}{4}\right)\sin\left(\frac{\pi}{L}z\right)\sin\left(\frac{2\pi}{L}z\right)dz = 2V_0. \end{aligned} \tag{9.304}$$

Following this procedure, we can obtain the remaining terms:

$$V = 2V_0 \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}. \tag{9.305}$$

The diagonalization of this matrix yields a doubly degenerate eigenvalue and a nondegenerate eigenvalue,

$$E_1^{(1)} = E_2^{(1)} = 0, \qquad E_3^{(1)} = 6V_0, \tag{9.306}$$

which lead to the energies of the first excited state:

$$E_1 = E_2 = \frac{3\pi^2\hbar^2}{mL^2}, \qquad E_3 = \frac{3\pi^2\hbar^2}{mL^2} + 6V_0. \tag{9.307}$$

So the perturbation has only partially lifted the degeneracy of the first excited state.

Problem 9.4

Consider a hydrogen atom which is subject to two *weak* static fields: an electric field in the xy planes $\vec{\mathcal{E}} = \mathcal{E}(\hat{\imath} + \hat{\jmath})$ and a magnetic field along the z-axis $\vec{B} = B\vec{k}$, where $\mathcal{E}$ and B are constant. Neglecting the spin–orbit interaction, calculate the energy levels of the $n = 2$ states to first-order perturbation.

Solution

In the absence of any external field, and neglecting spin–orbit interactions, the energy of the $n = 2$ state is fourfold degenerate: four different states $|\, nlm\rangle = |\, 200\rangle, |\, 211\rangle, |\, 210\rangle$, and $|\, 21-1\rangle$

correspond to the same energy $E_2 = -\mathcal{R}/4$, where $\mathcal{R} = m_e e^4/(2\hbar^2) = 13.6\,\text{eV}$ is the Rydberg constant.

When the atom is placed in an external electric field $\vec{\mathcal{E}} = \mathcal{E}(\hat{i} + \hat{j})$, the energy of interaction between the electron's dipole moment ($\vec{d} = -e\vec{r}$) and $\vec{\mathcal{E}}$ is given by $-\vec{d}\cdot\vec{\mathcal{E}} = e\mathcal{E}(x + y) = e\mathcal{E}r\sin\theta(\cos\phi + \sin\phi)$. On the other hand, when subjecting the atom to an external magnetic field $\vec{B} = B\vec{k}$, the linear momentum of the electron becomes $\vec{p} \longrightarrow (\vec{p} - e\vec{A}/c)$, where $\vec{A}$ is the vector potential corresponding to $\vec{B}$. So when subjecting a hydrogen atom to both $\vec{\mathcal{E}}$ and $\vec{B}$, its Hamiltonian is given by

$$\begin{aligned}\hat{H} &= \frac{1}{2\mu}\left(\vec{p} - \frac{e}{c}\vec{A}\right)^2 - \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{r} + e\mathcal{E}r\sin\theta(\cos\phi + \sin\phi) \\ &= \frac{\vec{p}^2}{2\mu} - \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{r} - \frac{e}{2\mu c}\vec{B}\cdot\vec{L} + \frac{e}{2\mu c}A^2 + e\mathcal{E}r\sin\theta(\cos\phi + \sin\phi).\end{aligned} \tag{9.308}$$

Since the magnetic field is weak, we can ignore the term $eA^2/(2\mu c)$; hence we can write $\hat{H}$ as $\hat{H} = \hat{H}_0 + \hat{H}_p$, where $\hat{H}_0$ is the Hamiltonian of an unperturbed hydrogen atom, while $\hat{H}_p$ can be treated as a perturbation:

$$\hat{H}_0 = \frac{\vec{p}^2}{2\mu} - \left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{r}, \qquad \hat{H}_p = -\frac{eB}{2\mu c}\hat{L}_z + e\mathcal{E}r\sin\theta(\cos\phi + \sin\phi). \tag{9.309}$$

To calculate the energy levels of the $n = 2$ state, we need to use degenerate perturbation theory, since the $n = 2$ state is fourfold degenerate; for this, we need to diagonalize the matrix

$$\begin{pmatrix} \langle 1|\hat{H}_p|1\rangle & \langle 1|\hat{H}_p|2\rangle & \langle 1|\hat{H}_p|3\rangle & \langle 1|\hat{H}_p|4\rangle \\ \langle 2|\hat{H}_p|1\rangle & \langle 2|\hat{H}_p|2\rangle & \langle 2|\hat{H}_p|3\rangle & \langle 2|\hat{H}_p|4\rangle \\ \langle 3|\hat{H}_p|1\rangle & \langle 3|\hat{H}_p|2\rangle & \langle 3|\hat{H}_p|3\rangle & \langle 3|\hat{H}_p|4\rangle \\ \langle 4|\hat{H}_p|1\rangle & \langle 4|\hat{H}_p|2\rangle & \langle 4|\hat{H}_p|3\rangle & \langle 4|\hat{H}_p|4\rangle \end{pmatrix}, \tag{9.310}$$

where $|1\rangle = |200\rangle$, $|2\rangle = |211\rangle$, $|3\rangle = |210\rangle$, and $|4\rangle = |21-1\rangle$. We therefore need to calculate the term

$$\langle 2l'm'|\hat{H}_p|2lm\rangle = -\frac{eB}{2\mu c}m\hbar\delta_{l',l}\delta_{m',m} + e\mathcal{E}\langle 2l'm'|r\sin\theta(\cos\phi + \sin\phi)|2lm\rangle. \tag{9.311}$$

Since $x = r\sin\theta\cos\phi$ and $y = r\sin\theta\sin\phi$ are both odd, the only terms that survive among $\langle 2l'm'|x|2lm\rangle$ and $\langle 2l'm'|y|2lm\rangle$ are $\langle 200|x|21\pm1\rangle$, $\langle 200|y|21\pm1\rangle$, and their complex conjugates. That is, x and y can couple only states of different parities ($l' - l = \pm1$) and whose azimuthal quantum numbers satisfy this condition: $m' - m = \pm1$. So we need to calculate only

$$\langle 200|x|21\pm1\rangle = \int_0^{+\infty} R_{20}^*(r)R_{21}(r)r^3dr \int Y_{00}^*(\Omega)\sin\theta\cos\phi Y_{1\pm1}(\Omega)d\Omega, \tag{9.312}$$

$$\langle 200|y|21\pm1\rangle = \int_0^{+\infty} R_{20}^*(r)R_{21}(r)r^3dr \int Y_{00}^*(\Omega)\sin\theta\sin\phi Y_{1\pm1}(\Omega)d\Omega, \tag{9.313}$$

where

$$\int_0^{+\infty} R_{20}^*(r)R_{21}(r)r^3dr = -3\sqrt{3}a_0; \tag{9.314}$$

a_0 is the Bohr radius, $a_0 = \hbar^2/(m_e e^2)$. Using the relations

$$\sin\theta\cos\phi = \sqrt{\frac{2\pi}{3}}\left[Y_{1-1}(\Omega) - Y_{11}(\Omega)\right], \qquad \sin\theta\sin\phi = i\sqrt{\frac{2\pi}{3}}\left[Y_{1-1}(\Omega) + Y_{11}(\Omega)\right], \tag{9.315}$$

along with

$$\int Y^*_{l'm'}(\Omega)Y_{lm}(\Omega)d\Omega = \delta_{l',l}\delta_{m',m}, \tag{9.316}$$

we obtain

$$\int Y^*_{00}(\Omega)\sin\theta\cos\phi Y_{11}(\Omega)d\Omega = \frac{1}{\sqrt{4\pi}}\int \sin\theta\cos\phi Y_{11}(\Omega)d\Omega = \frac{1}{\sqrt{6}}\int Y_{1-1}(\Omega)Y_{11}(\Omega)d\Omega$$
$$= -\frac{1}{\sqrt{6}}, \tag{9.317}$$

$$\int Y^*_{00}(\Omega)\sin\theta\sin\phi Y_{11}(\Omega)d\Omega = \frac{i}{\sqrt{6}}\int Y_{1-1}(\Omega)Y_{11}(\Omega)d\Omega = -\frac{i}{\sqrt{6}}. \tag{9.318}$$

Similarly, we have

$$\int Y^*_{00}(\Omega)\sin\theta\cos\phi Y_{1-1}(\Omega)d\Omega = \frac{1}{\sqrt{6}}. \tag{9.319}$$

$$\int Y^*_{00}(\Omega)\sin\theta\sin\phi Y_{1-1}(\Omega)d\Omega = -\frac{i}{\sqrt{6}}. \tag{9.320}$$

Now, substituting (9.314), (9.317), and (9.319) into (9.312), we end up with

$$\langle 200 \mid x \mid 21\pm 1\rangle = \pm\frac{3}{\sqrt{2}}a_0, \qquad \langle 200 \mid y \mid 21\pm 1\rangle = \frac{3i}{\sqrt{2}}a_0; \tag{9.321}$$

hence

$$\langle 21\pm 1|x \mid 200\rangle = \pm\frac{3}{\sqrt{2}}a_0, \qquad \langle 21\pm 1|y \mid 200\rangle = -\frac{3i}{\sqrt{2}}a_0. \tag{9.322}$$

The matrix (9.310) thus becomes

$$\begin{pmatrix} 0 & \alpha + i\alpha & 0 & -\alpha + i\alpha \\ \alpha - i\alpha & -\beta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\alpha - i\alpha & 0 & 0 & \beta \end{pmatrix}, \tag{9.323}$$

where α and β stand for $\alpha = 3e\mathcal{E}a_0/\sqrt{2}$ and $\beta = e\hbar B/(2\mu c)$.

The diagonalization of (9.323) yields the following eigenvalues:

$$\lambda_1 = -\sqrt{\frac{e^2\hbar^2B^2}{4\mu^2c^2} + 18e^2\mathcal{E}^2a_0^2}, \quad \lambda_2 = \lambda_3 = 0, \quad \lambda_4 = \sqrt{\frac{e^2\hbar^2B^2}{4\mu^2c^2} + 18e^2\mathcal{E}^2a_0^2}. \tag{9.324}$$

Finally, the energy levels of the $n = 2$ states are given to first-order approximation by

$$E^{(1)}_{2_1} = -\frac{\mathcal{R}}{4} - \sqrt{\frac{e^2\hbar^2B^2}{4\mu^2c^2} + 18e^2\mathcal{E}^2a_0^2}, \qquad E^{(1)}_{2_2} = -\frac{\mathcal{R}}{4}, \tag{9.325}$$

$$E_{2_3}^{(1)} = -\frac{\mathcal{R}}{4}, \qquad E_{2_4}^{(1)} = -\frac{\mathcal{R}}{4} + \sqrt{\frac{e^2\hbar^2B^2}{4\mu^2c^2} + 18e^2\mathcal{E}^2a_0^2}. \tag{9.326}$$

So the external electric and magnetic fields have lifted the degeneracy of the $n = 2$ level only partially.

Problem 9.5
A system, with an unperturbed Hamiltonian $\hat{H}_0$, is subject to a perturbation $\hat{H}_1$ with

$$\hat{H}_0 = E_0\begin{pmatrix} 15 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \qquad \hat{H}_1 = \frac{E_0}{100}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(a) Find the eigenstates of the unperturbed Hamiltonian $\hat{H}_0$ as well as the exact eigenvalues of the total Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}_p$.

(b) Find the eigenenergies of $\hat{H}$ to first-order perturbation. Compare them with the exact values obtained in (a).

Solution

(a) First, a diagonalization of $\hat{H}_0$ yields the eigenstates

$$|\phi_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |\phi_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\phi_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |\phi_4\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{9.327}$$

The values of the unperturbed energies are given by a nondegenerate value $E_1^{(0)} = 15E_0$ and a *threefold degenerate* value $E_2^{(0)} = E_3^{(0)} = E_4^{(0)} = 3E_0$.

The *exact* eigenvalues of $\hat{H}$ can be obtained by diagonalizing $\hat{H}$. Adopting the notation $\lambda = 1/100$, we can write the secular equation as

$$\begin{vmatrix} 15E_0 - E & 0 & 0 & 0 \\ 0 & 3E_0 - E & \lambda E_0 & 0 \\ 0 & \lambda E_0 & 3E_0 - E & 0 \\ 0 & 0 & 0 & 3E_0 - E \end{vmatrix} = 0 \tag{9.328}$$

or

$$(15E_0 - E)(3E_0 - E)\left[(3E_0 - E)^2 - \lambda^2 E_0^2\right] = 0, \tag{9.329}$$

which in turn leads to the *exact* values of the eigenenergies:

$$E_1 = 15E_0, \quad E_2 = 3E_0, \quad E_3 = (3 - \lambda)E_0, \quad E_4 = (3 + \lambda)E_0. \tag{9.330}$$

(b) To calculate the energy eigenvalues of $\hat{H}$ to first-order degenerate perturbation, and since $\hat{H}_0$ has one nondegenerate eigenvalue, $15E_0$, and a threefold degenerate eigenvalue, $3E_0$,

we need to make use of both nondegenerate and degenerate perturbative treatments. First, let us focus on the nondegenerate state; its energy is given by

$$\begin{aligned} E_1 &= 15E_0 + \langle \phi_1 \mid \hat{H}_1 \mid \phi_1 \rangle \\ &= 15E_0 + \frac{E_0}{100}(1\,0\,0\,0)\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda E_0 & 0 \\ 0 & \lambda E_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ &= 15E_0. \end{aligned} \tag{9.331}$$

This is identical with the *exact* eigenvalue (9.330) obtained in (a).

Second, to find the degenerate states, we need to diagonalize the matrix

$$V = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{pmatrix}, \tag{9.332}$$

where

$$V_{11} = \langle \phi_2 \mid \hat{H}_p \mid \phi_2 \rangle = (0\,1\,0\,0)\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda E_0 & 0 \\ 0 & \lambda E_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = 0, \tag{9.333}$$

$$V_{12} = \langle \phi_2 \mid \hat{H}_p \mid \phi_3 \rangle = (0\,1\,0\,0)\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda E_0 & 0 \\ 0 & \lambda E_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \lambda E_0, \tag{9.334}$$

$$V_{13} = \langle \phi_2 \mid \hat{H}_p \mid \phi_4 \rangle = (0\,1\,0\,0)\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda E_0 & 0 \\ 0 & \lambda E_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = 0. \tag{9.335}$$

Similarly, we can show that

$$V_{21} = \langle \phi_3 \mid \hat{H}_p \mid \phi_2 \rangle = \lambda E_0, \quad V_{22} = \langle \phi_3 \mid \hat{H}_p \mid \phi_3 \rangle = 0, \quad V_{23} = \langle \phi_3 \mid \hat{H}_p \mid \phi_4 \rangle = 0, \tag{9.336}$$

$$V_{31} = \langle \phi_4 \mid \hat{H}_p \mid \phi_2 \rangle = 0, \quad V_{32} = \langle \phi_4 \mid \hat{H}_p \mid \phi_3 \rangle = 0, \quad V_{33} = \langle \phi_4 \mid \hat{H}_p \mid \phi_4 \rangle = 0. \tag{9.337}$$

So the diagonalization of

$$V = \begin{pmatrix} 0 & \lambda E_0 & 0 \\ \lambda E_0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{9.338}$$

leads to the corrections $E_2^{(1)} = 0$, $E_3^{(1)} = \lambda E_0$, and $E_4^{(1)} = -\lambda E_0$. Thus, the energy eigenvalues to first-order degenerate perturbation are

$$E_2 = E_2^{(0)} + E_2^{(1)} = 3E_0, \qquad E_3 = E_3^{(0)} + E_3^{(1)} = (3 - \lambda)E_0, \tag{9.339}$$

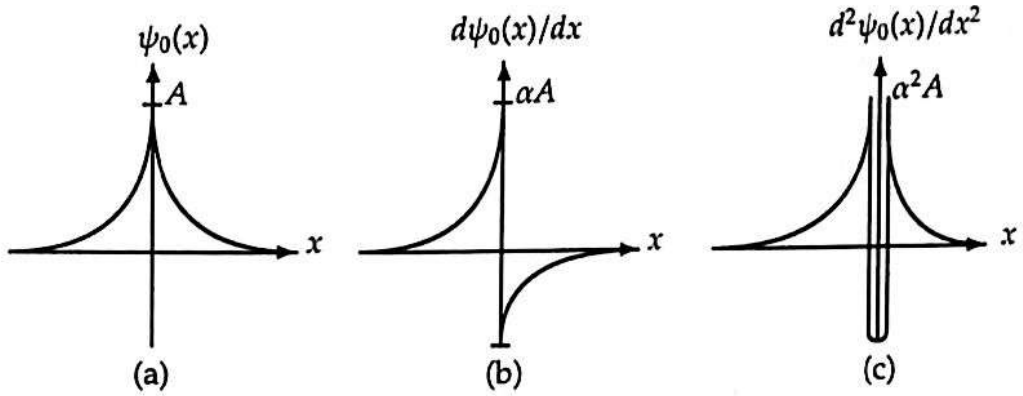


Figure 9.9: Shapes of $\psi_0(x) = Ae^{-\alpha|x|}$, $d\psi_0(x)/dx$, and $d^2\psi_0(x)/dx^2$.

$$E_4 = E_4^{(0)} + E_4^{(1)} = (3 + \lambda)E_0. \tag{9.340}$$

These are indeed identical with the *exact* eigenenergies (9.330) obtained in (a).

Problem 9.6

Use the variational method to estimate the energy of the ground state of a one-dimensional harmonic oscillator by making use of the following two trial functions:

(a) $\psi_0(x, \alpha) = Ae^{-\alpha|x|}$, (b) $\psi_0(x, \alpha) = A/(x^2 + \alpha)$,

where α is a positive real number and where A is the normalization constant.

Solution

(a) This wave function, whose shape is displayed in Figure 9.9a, is quite different from a Gaussian: it has a cusp at $x = 0$; hence its first derivative is discontinuous at $x = 0$.

The normalization constant A can be calculated at once:

$$\langle \psi_0 \mid \psi_0 \rangle = A^2 \int_{-\infty}^{0} e^{2\alpha x} dx + A^2 \int_0^{\infty} e^{-2\alpha x} dx = 2A^2 \int_0^{\infty} e^{-2\alpha x} dx = \frac{A^2}{\alpha}; \tag{9.341}$$

hence $A = \sqrt{\alpha}$. To find $E_0(\alpha)$ we need to calculate the potential and the kinetic terms. Using the integral $\int_0^{+\infty} x^n e^{-ax} dx = n!/a^{n+1}$ we can easily calculate the potential term:

$$\langle \psi_0 | V(x) \mid \psi_0 \rangle = \frac{1}{2} m\omega^2 A^2 \int_{-\infty}^{+\infty} x^2 e^{-2\alpha|x|} dx = m\omega^2 A^2 \int_0^{+\infty} x^2 e^{-2\alpha x} dx = \frac{m\omega^2}{4\alpha^2}. \tag{9.342}$$

But the kinetic energy term $-(\hbar^2/2m)\langle \psi_0 | d^2/dx^2 | \psi_0 \rangle$ is quite tricky to calculate. Since the first derivative of $\psi_0(x)$ is discontinuous at $x = 0$, a careless, straightforward calculation of $\langle \psi_0 | d^2/dx^2 | \psi_0 \rangle$, which makes use of (9.134), leads to a negative kinetic energy:

$$\begin{aligned} -\frac{\hbar^2}{2m} \left\langle \psi_0 \left| \frac{d^2}{dx^2} \right| \psi_0 \right\rangle &= -\frac{\hbar^2}{2m} A^2 \int_{-\infty}^{+\infty} e^{-\alpha|x|} \frac{d^2 e^{-\alpha|x|}}{dx^2} dx \\ &= -\frac{\hbar^2}{m} A^2 \int_0^{+\infty} e^{-\alpha x} \frac{d^2 e^{-\alpha x}}{dx^2} dx \\ &= -\frac{\hbar^2 \alpha^2}{m} A^2 \int_0^{+\infty} e^{-2\alpha x} dx = -\frac{\hbar^2 \alpha^2}{2m}. \end{aligned} \tag{9.343}$$

So when the first derivative of the wave function is discontinuous, the correct way to calculate the kinetic energy term is by using (9.135):

$$-\frac{\hbar^2}{2m}\left\langle\psi_0\left|\frac{d^2}{dx^2}\right|\psi_0\right\rangle = \frac{\hbar^2}{2m}A^2\int_{-\infty}^{+\infty}\left|\frac{de^{-\alpha|x|}}{dx}\right|^2 dx = \frac{\hbar^2\alpha^2}{2m}A^2\int_{-\infty}^{+\infty}e^{-2\alpha|x|}dx$$
$$= \frac{\hbar^2\alpha^2}{2m}, \tag{9.344}$$

because $A^2\int_{-\infty}^{+\infty}e^{-2\alpha|x|}dx = 1$.

Why do expressions (9.343) and (9.344) yield different results? The reason is that the correct expression of $d^2e^{-\alpha|x|}/dx^2$ must involve a delta function (Figures 9.9a and 9.9b). That is, the correct form of $d\psi(x)/dx$ is given by

$$\frac{d\psi_0(x)}{dx} = A\frac{de^{-\alpha|x|}}{dx} = -\alpha\psi_0(x)\frac{d|x|}{dx} = -\alpha\psi_0(x)\begin{cases}-1, & x<0,\\ 1, & x>0.\end{cases} \tag{9.345}$$

or

$$\frac{d\psi_0(x)}{dx} = -\alpha\left[\Theta(x)-\Theta(-x)\right]\psi_0(x), \tag{9.346}$$

where $\Theta(x)$ is the Heaviside function

$$\Theta(x) = \begin{cases}0, & x<0,\\ 1, & x>0.\end{cases} \tag{9.347}$$

The second derivative of $\psi_0(x)$ therefore contains a delta function:

$$\frac{d^2\psi_0(x)}{dx^2} = \frac{d}{dx}\left\{-\alpha\left[\Theta(x)-\Theta(-x)\right]\psi_0(x)\right\}, \tag{9.348}$$

and since

$$\frac{d\Theta(x)}{dx} = \delta(x), \qquad \left[\Theta(x)-\Theta(-x)\right]^2 = 1, \tag{9.349}$$

and since $\delta(x) = \delta(-x)$, we have

$$\frac{d^2\psi_0(x)}{dx^2} = \alpha^2\left[\Theta(x)-\Theta(-x)\right]^2\psi_0(x) - \alpha\left[\delta(x)+\delta(-x)\right]\psi_0(x)$$
$$= \alpha^2\psi_0(x) - 2\alpha\psi_0(x)\delta(x). \tag{9.350}$$

So the substitution of (9.350) into (9.343) leads to the same (correct) expression as (9.344):

$$-\frac{\hbar^2}{2m}\left\langle\psi_0\left|\frac{d^2}{dx^2}\right|\psi_0\right\rangle = -\frac{\hbar^2}{2m}\int_{-\infty}^{+\infty}\psi_0^*(x)\frac{d^2\psi_0(x)}{dx^2}dx$$
$$= -\frac{\hbar^2}{2m}\int_{-\infty}^{+\infty}\psi_0^*(x)\left[\alpha^2\psi_0(x) - 2\alpha\psi_0(x)\delta(x)\right]dx$$
$$= -\frac{\hbar^2}{2m}\alpha^2 + \frac{\hbar^2}{m}\alpha\left|\psi_0(0)\right|^2 = -\frac{\hbar^2}{2m}\alpha^2 + \frac{\hbar^2}{m}\alpha^2$$
$$= \frac{\hbar^2}{2m}\alpha^2. \tag{9.351}$$

Now, adding (9.342) and (9.351), we get

$$E_0(\alpha) = \frac{\hbar^2}{2m}\alpha^2 + \frac{m\omega^2}{4\alpha^2}. \tag{9.352}$$

The minimization of $E_0(\alpha)$,

$$0 = \frac{\partial E_0(\alpha)}{\partial \alpha} = \frac{\hbar^2}{m}\alpha_0 - \frac{m\omega^2}{2\alpha_0^3}, \tag{9.353}$$

leads to $\alpha_0^2 = m\omega/(\sqrt{2}\hbar)$ which, when inserted into (9.352), leads to

$$E_0(\alpha_0) = \frac{\hbar^2}{2m}\frac{m\omega}{\sqrt{2}\hbar} + \frac{m\omega^2}{4}\frac{\sqrt{2}\hbar}{m\omega} = \frac{\hbar\omega}{\sqrt{2}} = 0.707\hbar\omega. \tag{9.354}$$

This inaccurate result was expected; it is due to the cusp at $x = 0$.

(b) We can show that the normalization constant A is given by $A = (4\alpha^3/\pi^2)^{1/4}$. Unlike $Ae^{-\alpha|x|}$, the first derivative of the trial $A/(1 + x^2)$ is continuous; hence we can use (9.134) to calculate the kinetic energy term. The ground state energy is given by

$$\begin{aligned} E_0(\alpha) &= \langle \psi_0(\alpha)|\hat{H}|\psi_0(\alpha)\rangle \\ &= A^2 \int_{-\infty}^{+\infty} \frac{1}{x^2+\alpha}\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2x^2\right)\frac{1}{x^2+\alpha}dx \\ &= -\frac{A^2\hbar^2}{2m}\int_{-\infty}^{+\infty}\frac{6x^2-2\alpha}{(x^2+\alpha)^4}dx + \frac{1}{2}m\omega^2A^2\int_{-\infty}^{+\infty}\frac{x^2}{(x^2+\alpha)^2}dx \\ &= \frac{\hbar^2}{4m\alpha} + \frac{1}{2}m\omega^2\alpha. \end{aligned} \tag{9.355}$$

The minimization of $E_0(\alpha)$ with respect to α (i.e., $\partial E(\alpha)/\partial\alpha = 0$) yields $\alpha_0 = \hbar/(\sqrt{2}\,m\omega)$ which, when inserted into (9.355), leads to

$$E_0(\alpha_0) = \frac{\hbar\omega}{\sqrt{2}}. \tag{9.356}$$

This energy, which is larger than the exact value $\hbar\omega/2$ by a factor of $\sqrt{2}$, is similar to that of part (a); this is a pure coincidence. The size of this error is due to the fact that the trial function $A/(x^2+\alpha)$ is not a good approximation to the exact wave function, which has a Gaussian form.

Problem 9.7

For a particle of mass m moving in a one-dimensional box with walls at $x = 0$ and $x = L$, use the variational method to estimate

(a) its ground state energy and

(b) its first excited state energy.

Solution

The exact solutions of this problem are known: $E_n^{exact} = \pi^2\hbar^2n^2/(2mL^2)$.

(a) The trial function for the ground state must vanish at the walls, it must have no nodes, and must be symmetric (i.e., even) with respect to $x = L/2$. These three requirements can be satisfied by the following parabolic trial function:

$$\psi_0(x) = x(L - x); \tag{9.357}$$

no scale parameter is needed here. Since no parameter is involved, we can calculate the energy directly (no variation is required): $E_0 = \langle\psi_0|\hat{H}\,|\,\psi_0\rangle/\langle\psi_0\,|\,\psi_0\rangle$, where

$$\langle\psi_0\,|\,\psi_0\rangle = \int_0^L \psi_0^2(x)\,dx = \int_0^L x^2(L^2 - 2Lx + x^2)\,dx = \frac{1}{30}L^5, \tag{9.358}$$

and

$$\langle\psi_0|\hat{H}\,|\,\psi_0\rangle = \frac{\hbar^2}{2m}\int_0^L \left(\frac{d\psi_0(x)}{dx}\right)^2 dx = \frac{\hbar^2}{2m}\int_0^L (L^2 - 4Lx + 4x^2)\,dx = \frac{\hbar^2 L^3}{6m}. \tag{9.359}$$

Thus, the ground state energy is given by

$$E_0^{VM} = \frac{\langle\psi_0|\hat{H}\,|\,\psi_0\rangle}{\langle\psi_0\,|\,\psi_0\rangle} = 10\frac{\hbar^2}{2mL^2}. \tag{9.360}$$

This is a very accurate result, for it is higher than the exact result by a mere 1%:

$$E_0^{VM} = \frac{10}{\pi^2}E^{exact}. \tag{9.361}$$

(b) The properties of the exact wave function of the first excited state are known: it has one node at $x = L/2$ and must be odd with respect to $x = L/2$; this last property makes it orthogonal to the ground state which is even about $L/2$. Let us try a polynomial function. Since the wave function vanishes at $x = 0, L/2$, and L, the trial function must be at least cubic. The following polynomial function satisfies all these conditions:

$$\psi_1(x) = x\left(x - \frac{L}{2}\right)(x - L). \tag{9.362}$$

Again, no scale parameter is needed.

To calculate E_1^{VM}, we need to find

$$\langle\psi_1\,|\,\psi_1\rangle = \int_0^L \psi_1^2(x)dx = \int_0^L x^2\left(x - \frac{L}{2}\right)^2(x - L)^2 dx = \frac{1}{840}L^7 \tag{9.363}$$

and

$$\begin{aligned}\langle\psi_1|\hat{H}\,|\,\psi_1\rangle &= \frac{\hbar^2}{2m}\int_0^L \left(\frac{d\psi_1(x)}{dx}\right)^2 dx = \frac{\hbar^2}{2m}\int_0^L \left(3x^2 - 3Lx + \frac{L^2}{2}\right)^2 dx \\ &= \frac{\hbar^2 L^5}{40m}.\end{aligned} \tag{9.364}$$

Dividing the previous two expressions, we obtain the energy of the first excited state:

$$E_1^{VM} = \frac{\langle\psi_1|\hat{H}\,|\,\psi_1\rangle}{\langle\psi_1\,|\,\psi_1\rangle} = 42\frac{\hbar^2}{2mL^2}. \tag{9.365}$$

This too is a very accurate result; since $E_1^{exact} = (2\pi)^2\hbar^2/(2mL^2)$ we can write E_1^{VM} as $E_1^{VM} = 42E_1^{exact}/(2\pi)^2$; hence E_1^{VM} is higher than E_1^{exact} by 6%.

Problem 9.8
Consider an infinite, one-dimensional potential well of length L, with walls at $x = 0$ and $x = L$, that is modified at the bottom by a perturbation $V_p(x)$:

$$V(x) = \begin{cases} 0, & 0 < x < L, \\ \infty, & \text{elsewhere,} \end{cases} \qquad V_p(x) = \begin{cases} V_0, & 0 \le x \le L/2, \\ 0, & \text{elsewhere,} \end{cases}$$

where $V_0 \ll 1$.

(a) Using first-order perturbation theory, calculate the energy E_n.

(b) Calculate the energy E_n in the WKB approximation. Compare this energy with the expression obtained in (a).

Solution
The exact energy E_n^{exact} and wave function $\phi_n(x)$ for a potential well are given by

$$E_n^{exact} = \frac{\pi^2\hbar^2}{2mL^2}n^2, \qquad \phi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right).$$

(a) Since the first-order correction to the energy caused by the perturbation $V_p(x)$ is given by

$$\begin{aligned} E_n^{(1)} = \langle \phi_n \mid V_p \mid \phi_n \rangle &= \frac{2}{L}V_0 \int_0^{L/2} \sin^2\left(\frac{n\pi x}{L}\right)dx \\ &= \frac{1}{L}V_0 \int_0^{L/2} \left[1 - \cos\left(\frac{2n\pi x}{L}\right)\right]dx = \frac{V_0}{2}; \end{aligned} \tag{9.366}$$

hence the energy is given to first-order perturbation by

$$E_n^{PT} = \frac{\pi^2\hbar^2}{2mL^2}n^2 + \frac{V_0}{2}. \tag{9.367}$$

(b) Since this potential has two rigid walls, the energy within the WKB approximation needs to be extracted from the quantization condition $\int_0^L p(E_n, x)\,dx = n\pi\hbar$, where

$$\begin{aligned} \int_0^L p(E_n, x)\,dx &= \sqrt{2m(E_n - V_0)}\int_0^{L/2} dx + \sqrt{2mE_n}\int_{L/2}^{L} dx \\ &= \frac{L}{2}\sqrt{2m}\left(\sqrt{E_n - V_0} + \sqrt{E_n}\right); \end{aligned} \tag{9.368}$$

hence $L\sqrt{2m}\left(\sqrt{E_n - V_0} + \sqrt{E_n}\right) = 2n\pi\hbar$ or

$$\sqrt{E_n - V_0} + \sqrt{E_n} = \frac{2n\pi\hbar}{L\sqrt{2m}}. \tag{9.369}$$

Squaring both sides of this equation and using the notation $\alpha_n = 2n^2\pi^2\hbar^2/(mL^2)$, we have

$$2\sqrt{E_n(E_n - V_0)} = \alpha_n - 2E_n + V_0. \tag{9.370}$$

Squaring both sides of this equation, we obtain

$$4E_n^2 - 4E_n V_0 = \alpha_n^2 + 4E_n^2 + V_0^2 - 4\alpha_n E_n + 2\alpha_n V_0 - 4E_n V_0, \tag{9.371}$$

which, solving for E_n, leads to

$$E_n = \frac{\alpha_n}{4} + \frac{V_0}{2} + \frac{V_0^2}{4\alpha_n} \tag{9.372}$$

or

$$E_n^{WKB} = \frac{\pi^2\hbar^2}{2mL^2}n^2 + \frac{V_0}{2} + \frac{mL^2V_0^2}{8\pi^2\hbar^2}\frac{1}{n^2}. \tag{9.373}$$

When $n \gg 1$, and since V_0 is very small, the WKB energy relation (9.373) gives back the expression (9.367) that was derived from a first-order perturbative treatment:

$$E_n^{WKB} \simeq E_n^{PT} = \frac{\pi^2\hbar^2}{2mL^2}n^2 + \frac{V_0}{2}.$$

Problem 9.9

Consider a particle of mass m that is bouncing vertically and elastically on a reflecting hard floor where $V(z) = \begin{cases} mgz, & z > 0, \\ +\infty, & z \leq 0, \end{cases}$ and g is the gravitational constant.

(a) Use the variational method to estimate the ground state energy of this particle.

(b) Use the WKB method to estimate the ground state energy of this particle.

(c) Compare the results of (a) and (b) with the exact ground state energy.

Solution

(a) The ground state wave function of this particle has no nodes and must vanish at $z = 0$ and be finite as $z \to +\infty$. The following trial function satisfies these conditions:

$$\psi_0(z, \alpha) = Aze^{-\alpha z}, \tag{9.374}$$

where α is a parameter and A is the normalization constant. We can show that $A = 2\alpha^{3/2}$ and hence

$$\psi_0(z, \alpha) = 2\sqrt{\alpha^3}ze^{-\alpha z}. \tag{9.375}$$

The energy is given by

$$\begin{aligned} E_0^{VM}(\alpha) &= 4\alpha^3 \int_0^{+\infty} ze^{-\alpha z}\left[-\frac{\hbar^2}{2m}\frac{d^2}{dz^2} + mgz\right]ze^{-\alpha z}\,dz \\ &= 4\alpha^3\frac{\hbar^2}{2m}\int_0^{+\infty}\left(2\alpha z - \alpha^2 z^2\right)e^{-2\alpha z}\,dz + 4\alpha^3 mg\int_0^{+\infty} z^3 e^{-2\alpha z}\,dz \\ &= 2\alpha^3\frac{\hbar^2}{m}\left(\frac{1}{2\alpha} - \frac{1}{4\alpha}\right) + 4mg\alpha^3\frac{3}{8\alpha^4}, \end{aligned} \tag{9.376}$$

or

$$E_0^{VM}(\alpha) = \frac{\hbar^2}{2m}\alpha^2 + \frac{3}{2\alpha}mg. \tag{9.377}$$

The minimization of $E_0(\alpha)$ yields $\alpha_0 = (3m^2g/2\hbar^2)^{1/3}$ and hence

$$E_0^{VM}(\alpha_0) = \frac{3}{2}\left(\frac{9}{2}\right)^{1/3}\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3}. \tag{9.378}$$

(b) Since this potential has one rigid wall at $x = 0$, the correct quantization rule is given by (9.235): $\int_0^{E/mg} p\,dz = (n + \frac{3}{4})\pi\hbar$; the turning point occurs at $E = mgz$ and hence $z = E/mg$. Now, since $E = p^2/2m + mgz$ we have $p(E,z) = \sqrt{2mE}\sqrt{1 - mgz/E}$, and therefore

$$\int_0^{E/mg} p(E_n,z)\,dz = \sqrt{2mE}\int_0^{E/mg}\sqrt{1-\frac{mg}{E}z}\,dz = \sqrt{2mE}\frac{2E}{3mg} = \sqrt{\frac{8E^3}{9mg^2}}. \tag{9.379}$$

Inserting this relation into the quantization condition $\int_0^{E/mg} p\,dz = (n + \frac{3}{4})\pi\hbar$ gives

$$\sqrt{\frac{8E^3}{9mg^2}} = \left(n + \frac{3}{4}\right)\pi\hbar, \tag{9.380}$$

and we obtain the WKB approximation for the energy:

$$E_n^{WKB} = \left[\frac{9\pi^2}{8}mg^2\pi^2\hbar^2\left(n+\frac{3}{4}\right)^2\right]^{1/3}. \tag{9.381}$$

Hence the ground state energy is given by

$$E_0^{WKB} = \frac{3}{4}\left(3\pi^2\right)^{1/3}\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3}. \tag{9.382}$$

(c) Recall that the *exact* ground state energy, calculated in Problem 4.18, page 304, for a particle of mass m moving in the potential $V(z) = mgz$ is given by

$$E_0^{exact} = 2.338\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3}. \tag{9.383}$$

Combining this relation with (9.378) and (9.382), we see that the variational method overestimates the energy by a 5.9% error, while the WKB method underestimates it by a 0.8% error:

$$E_0^{VM} = \frac{3}{2}\left(\frac{9}{2}\right)^{1/3}\frac{E_0^{exact}}{2.338} \simeq 1.059E_0^{exact}, \tag{9.384}$$

$$E_0^{WKB} = \frac{3}{4}\left(3\pi^2\right)^{1/3}\frac{E_0^{exact}}{2.338} \simeq 0.992E_0^{exact}. \tag{9.385}$$

The variational method has given a reasonably accurate result because we succeeded quite well in selecting the trial function. As for the WKB method, it has given a very accurate result because we have used the correct quantization rule (9.235). Had we used the quantization rule (9.221), which contains a factor of $\frac{1}{2}$ instead of $\frac{1}{4}$ in (9.235), the WKB method would have given a very inaccurate result with a 24.3% error, i.e., $E_0^{WKB} \simeq 0.757E_0^{exact}$.

Problem 9.10
Using first-order perturbation theory, and ignoring the spin of the electron, calculate the energy of the 2p level of a hydrogen atom when placed in a *weak* quadrupole field whose principal axes are along the xyz axes: $\hat{H}_p = \sum_{\mu=-2}^{\mu=2} Q_\mu r^2 Y_{2\mu}(\Omega)$, where Q_μ are real numbers, with $Q_{-1} = Q_1 = 0$ and $Q_{-2} = Q_2$, and $Y_{2\mu}(\Omega)$ are spherical harmonics.

Solution
In the absence of the field, the energy levels of the $|\,2,\,1,\,m\rangle$ states are threefold degenerate: $|\,2,\,1,\,-1\rangle$, $|\,2,\,1,\,0\rangle$, and $|\,2,\,1,\,1\rangle$, and hence correspond to the same energy $E_2 = -\mathcal{R}/4$, where $\mathcal{R} = 13.6\,\text{eV}$ is the Rydberg constant.

When the quadrupole field is turned on, and since Q_{-2}, Q_0, and Q_2 are small, we can treat the quadrupole interaction $\hat{H}_p = Q_{-2}r^2Y_{2-2}(\Omega) + Q_0 r^2 Y_{20}(\Omega) + Q_2 r^2 Y_{22}(\Omega)$ as a perturbation. To calculate the p split, we need to use degenerate perturbation theory, which, in a first step, requires calculating the matrix

$$\begin{pmatrix} \langle 2,\,1,\,-1|\hat{H}_p\,|\,2,\,1,\,-1\rangle & \langle 2,\,1,\,-1|\hat{H}_p\,|\,2,\,1,\,0\rangle & \langle 2,\,1,\,-1|\hat{H}_p\,|\,2,\,1,\,1\rangle \\ \langle 2,\,1,\,0|\hat{H}_p\,|\,2,\,1,\,-1\rangle & \langle 2,\,1,\,0|\hat{H}_p\,|\,2,\,1,\,0\rangle & \langle 2,\,1,\,0|\hat{H}_p\,|\,2,\,1,\,1\rangle \\ \langle 2,\,1,\,1|\hat{H}_p\,|\,2,\,1,\,-1\rangle & \langle 2,\,1,\,1|\hat{H}_p\,|\,2,\,1,\,0\rangle & \langle 2,\,1,\,1|\hat{H}_p\,|\,2,\,1,\,1\rangle \end{pmatrix}, \tag{9.386}$$

where

$$\langle 2,\,1,\,m'|\hat{H}_p\,|\,2,\,1,\,m\rangle = \langle 2,\,1|r^2|2,\,1\rangle\langle 1,\,m'|Q_{-2}Y_{2-2} + Q_0Y_{20} + Q_2Y_{22}|1,\,m\rangle. \tag{9.387}$$

The radial part is easy to obtain (Chapter 6):

$$\langle n,\,l|r^2|n,\,l\rangle = \int_0^{+\infty} r^4\,|R_{nl}|^2\,dr = \frac{1}{2}n^2\left[5n^2 + 1 - 3l(l+1)\right]a_0^2; \tag{9.388}$$

hence

$$\langle 2,\,1|r^2|2,\,1\rangle = 30a_0^2. \tag{9.389}$$

As for the angular part, it can be inferred from the Wigner–Eckart theorem:

$$\langle l',\,m'|Y_{2\mu}|1,\,m\rangle = \langle l,2;\,m,\mu|l',\,m'\rangle\langle l'\,\|\,Y_2\,\|\,1\rangle; \tag{9.390}$$

the reduced matrix element $\langle l'\,\|\,Y_2\,\|\,1\rangle$ was calculated in Chapter 7: $\langle l'\,\|\,Y_2\,\|\,1\rangle = \sqrt{5/4\pi}\,\sqrt{(2l+1)/(2l'+1)}\langle l,2;\,0,0|l',\,0\rangle$ and hence

$$\langle l',\,m'|Y_{2\mu}|1,\,m\rangle = \sqrt{\frac{5}{4\pi}}\sqrt{\frac{2l+1}{2l'+1}}\langle l,2;\,0,0|l',\,0\rangle\langle l,2;\,m,\mu|l',\,m'\rangle. \tag{9.391}$$

Using the coefficients $\langle l,2;\,m,0|l,\,m\rangle = [3m^2 - l(l+1)]/\sqrt{l(2l-1)(l+1)(2l+3)}$ and $\langle l,2;\,m,\mp2,\,\pm2|l,\,m\rangle = \sqrt{\frac{3(l\pm m-1)(l\pm m)(l\mp m+1)(l\mp m+2)}{2l(2l-1)(l+1)(2l+3)}}$, we have

$$\langle 1,\,-1\,|\,Y_{2-2}\,|\,1,\,1\rangle = \langle 1,\,1\,|\,Y_{22}\,|\,1,\,-1\rangle = -\sqrt{\frac{3}{10\pi}}, \tag{9.392}$$

$$\langle 1,\,-1\,|\,Y_{20}\,|\,1,\,-1\rangle = \langle 1,\,1\,|\,Y_{20}\,|\,1,\,1\rangle = -\frac{1}{\sqrt{20\pi}}, \tag{9.393}$$

$$\langle 1,\,0\,|\,Y_{20}\,|\,1,\,0\rangle = \frac{1}{\sqrt{5\pi}}. \tag{9.394}$$

These expressions can also be obtained from the following relations:

$$\int Y^*_{lm}(\Omega)Y_{20}(\Omega)Y_{lm}(\Omega)\,d\Omega = \sqrt{\frac{5}{4\pi}}\frac{l(l+1)-3m^2}{(2l-1)(2l+3)}, \tag{9.395}$$

$$\int Y^*_{lm+2}(\Omega)Y_{22}(\Omega)Y_{lm}(\Omega)d\Omega = \int Y^*_{lm}(\Omega)Y_{2-2}(\Omega)Y_{lm+2}(\Omega)d\Omega$$
$$= \sqrt{\frac{15}{8\pi}}\frac{\sqrt{(l-m-1)(l-m)(l+m+1)(l+m+2)}}{(2l-1)(2l+3)}. \tag{9.396}$$

Combining (9.387) to (9.394) we can write the matrix (9.386) as

$$30a_0^2\begin{pmatrix} -\frac{Q_0}{\sqrt{20\pi}} & 0 & -Q_2\sqrt{\frac{3}{10\pi}} \\ 0 & \frac{Q_0}{\sqrt{5\pi}} & 0 \\ -Q_2\sqrt{\frac{3}{10\pi}} & 0 & -\frac{Q_0}{\sqrt{20\pi}} \end{pmatrix}. \tag{9.397}$$

The diagonalization of this matrix leads to the following eigenvalues:

$$E_1^{(1)} = -30\frac{a_0^2}{\sqrt{10\pi}}\left(\frac{Q_0}{\sqrt{2}} + Q_2\sqrt{3}\right), \tag{9.398}$$

$$E_2^{(1)} = 30\frac{Q_0a_0^2}{\sqrt{5\pi}}, \tag{9.399}$$

$$E_3^{(1)} = 30\frac{a_0^2}{\sqrt{10\pi}}\left(-\frac{Q_0}{\sqrt{2}} + Q_2\sqrt{3}\right). \tag{9.400}$$

Thus, to first-order perturbation theory, the energies of the p level are given by

$$E_{21} = -\frac{\mathcal{R}}{4} - 30\frac{a_0^2}{\sqrt{10\pi}}\left(\frac{Q_0}{\sqrt{2}} + Q_2\sqrt{3}\right), \tag{9.401}$$

$$E_{22} = -\frac{\mathcal{R}}{4} + 30\frac{Q_0a_0^2}{\sqrt{5\pi}}, \tag{9.402}$$

$$E_{23} = -\frac{\mathcal{R}}{4} + 30\frac{a_0^2}{\sqrt{10\pi}}\left(-\frac{Q_0}{\sqrt{2}} + Q_2\sqrt{3}\right). \tag{9.403}$$

So the quadrupole interaction has lifted all the degeneracies of the p level.

Problem 9.11

Two protons, located on the z-axis and separated by a distance d (i.e., $\vec{r} = d\hat{k}$), are subject to a z-oriented magnetic field $\vec{B} = B\hat{k}$.

(a) Ignoring all interactions between the two protons, find the energy levels and stationary states of this system.

(b) Treating the dipole–dipole magnetic interaction energy between the protons,

$$\hat{H}_p = \frac{1}{r^3}\left[\vec{\mu}_1 \cdot \vec{\mu}_2 - 3\frac{(\vec{\mu}_1 \cdot \vec{r})(\vec{\mu}_2 \cdot \vec{r})}{r^2}\right],$$

as a perturbation, calculate the energy using first-order perturbation theory.

Solution

(a) Since the magnetic moments of the protons are $\vec{\mu}_1 = 2\mu_0\vec{S}_1/\hbar$ and $\vec{\mu}_2 = 2\mu_0\vec{S}_2/\hbar$ where $\mu_0 = \hbar e/(2M_p c)$ is the proton magnetic moment, the Hamiltonian of the two-proton system, ignoring all the interactions between the two protons, is due to the interaction of the magnetic moments of the protons with the external magnetic field:

$$\hat{H}_0 = -(\vec{\mu}_1 + \vec{\mu}_2)\cdot\vec{B} = -\frac{2\mu_0}{\hbar}(\vec{S}_1 + \vec{S}_2)\cdot\vec{B} = -\frac{2\mu_0 B}{\hbar}\hat{S}_z. \tag{9.404}$$

As shown in Chapter 7, the eigenstates of a system consisting of two spin-$\frac{1}{2}$ particles are a triplet state and singlet state; the stationary eigenstates of $\hat{H}_0$ are therefore given by

$$|\chi_1\rangle = |1, 1\rangle = \left|\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2}\right\rangle, \tag{9.405}$$

$$|\chi_2\rangle = |1, -1\rangle = \left|\frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2}\right\rangle, \tag{9.406}$$

$$|\chi_3\rangle = |1, 0\rangle = \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\right\rangle + \left|\frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2}\right\rangle\right], \tag{9.407}$$

$$|\chi_4\rangle = |0, 0\rangle = \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\right\rangle - \left|\frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2}\right\rangle\right]. \tag{9.408}$$

The eigenenergies of $|\chi_1\rangle$, $|\chi_2\rangle$, $|\chi_3\rangle$, and $|\chi_4\rangle$ are respectively

$$E_1^{(0)} = -2\mu_0 B, \qquad E_2^{(0)} = 2\mu_0 B, \qquad E_3^{(0)} = E_4^{(0)} = 0. \tag{9.409}$$

So $|\chi_3\rangle$ and $|\chi_4\rangle$ are (doubly) degenerate, whereas $|\chi_1\rangle$ and $|\chi_2\rangle$ are not.

(b) To calculate the energy to first order, we need to calculate the matrix elements of $\hat{H}_p$: $\hat{H}_{p_{ij}} = \langle\chi_i|\hat{H}_p|\chi_j\rangle$, with $i, j = 1, 2, 3, 4$. For this, since $\vec{r} = d\hat{k}$, we have $\vec{\mu}_1 \cdot \vec{r} = 2\mu_0 d\hat{S}_{1_z}/\hbar$ and $\vec{\mu}_2 \cdot \vec{r} = 2\mu_0 d\hat{S}_{2_z}/\hbar$. Thus, we can write $\hat{H}_p$ as

$$\hat{H}_p = \frac{1}{r^3}\left[\vec{\mu}_1 \cdot \vec{\mu}_2 - 3\frac{(\vec{\mu}_1 \cdot \vec{r})(\vec{\mu}_2 \cdot \vec{r})}{r^2}\right] = \frac{4\mu_0^2}{d^3\hbar^2}\left[\hat{\vec{S}}_1 \cdot \hat{\vec{S}}_2 - 3\hat{S}_{1_z}\hat{S}_{2_z}\right]. \tag{9.410}$$

Using the relations

$$\begin{aligned} 2\vec{S}_1 \cdot \vec{S}_2 \,|\, S, S_z\rangle &= \left[\left(\vec{S}_1 + \vec{S}_2\right)^2 - \vec{S}_1^2 - \vec{S}_1^2\right] |\, S, S_z\rangle \\ &= \hbar^2\left[S(S+1) - S_1(S_1+1) - S_2(S_2+1)\right] |\, S, S_z\rangle \\ &= \hbar^2\left[S(S+1) - \frac{3}{2}\right] |\, S, S_z\rangle, \end{aligned} \tag{9.411}$$

$$2\hat{S}_{1_z}\hat{S}_{2_z} \mid S,\ S_z\rangle = \left[\hat{S}_z{}^2 - \hat{S}_{1_z}{}^2\hat{S}_{2_z}{}^2\right] \mid S,\ S_z\rangle = \hbar^2\left(S_z^2 - \frac{1}{2}\right) \mid S,\ S_z\rangle, \tag{9.412}$$

along with (9.410), we can rewrite

$$\begin{aligned}\hat{H}_p \mid S,\ S_z\rangle &= \frac{2\mu_0^2}{d^3}\left[S(S+1) - \frac{3}{2} - 3\left(S_z^2 - \frac{1}{2}\right)\right] \mid S,\ S_z\rangle \\ &= \frac{2\mu_0^2}{d^3}\left[S(S+1) - 3S_z^2\right] \mid S,\ S_z\rangle. \end{aligned} \tag{9.413}$$

Since the values of S and S_z are given for the triplet state by $S = 1$, $S_z = -1, 0, 1$, and by $S = 0$, $S_z = 0$ for the singlet, the matrix elements of $\hat{H}_p$ are

$$E_1^{(1)} = \langle\chi_1 \mid \hat{H}_p \mid \chi_1\rangle = -\frac{2\mu_0^2}{d^3}, \qquad E_2^{(1)} = \langle\chi_2 \mid \hat{H}_p \mid \chi_2\rangle = -\frac{2\mu_0^2}{d^3}, \tag{9.414}$$

$$E_3^{(1)} = \langle\chi_3 \mid \hat{H}_p \mid \chi_3\rangle = \frac{4\mu_0^2}{d^3}, \qquad E_4^{(1)} = \langle\chi_4 \mid \hat{H}_p \mid \chi_4\rangle = 0. \tag{9.415}$$

All the other matrix elements of $\hat{H}_p$ are zero: $\langle\chi_i|\hat{H}_p|\chi_j\rangle = 0$ for $i \neq j$.

Finally, the energy levels of the two-proton system can be obtained at once from (9.409) along with (9.414) and (9.415):

$$E_1 = E_1^{(0)} + E_1^{(1)} = -2\mu_0 B - \frac{2\mu_0^2}{d^3}, \tag{9.416}$$

$$E_2 = E_2^{(0)} + E_2^{(1)} = 2\mu_0 B - \frac{2\mu_0^2}{d^3}, \tag{9.417}$$

$$E_3 = E_3^{(0)} + E_3^{(1)} = \frac{4\mu_0^2}{d^3}, \tag{9.418}$$

$$E_4 = E_4^{(0)} + E_4^{(1)} = 0. \tag{9.419}$$

So the dipole–dipole magnetic interaction has lifted the degeneracy of the energy levels in the two-proton system.

Problem 9.12

A spin $\frac{1}{2}$ particle of mass m, which is moving in an infinite, symmetric potential well $V(x)$ of length $2L$, is placed in an external *weak* magnetic field $\vec{B}$ with

$$V(x) = \begin{cases} 0, & -L \le x \le L, \\ \infty, & \text{otherwise}, \end{cases} \qquad \vec{B} = \begin{cases} -B\hat{z}, & -L \le x \le 0, \\ -B\hat{x}, & 0 \le x \le L. \end{cases}$$

Using first-order perturbation theory, calculate the energy of the nth excited state of this particle.

Solution

First, let us discuss the physics of this particle before placing it in a magnetic field. As seen in Chapter 4, the energy and wave function of a *spinless* particle of mass m moving in a symmetric

potential well of length $2L$ are

$$E_n = \frac{\hbar^2\pi^2}{8mL^2}n^2, \qquad \psi_n(x) = \frac{1}{\sqrt{L}}\begin{cases} \cos\left(\frac{n\pi x}{2L}\right), & n = 1,3,5,\ldots, \\ \sin\left(\frac{n\pi x}{2L}\right), & n = 2,4,6,\ldots. \end{cases} \tag{9.420}$$

When the spin of the particle is considered, its wave function is the product of a spacial part $\psi_n(x)$ and a spin part $|\chi_\pm\rangle$:

$$\psi_n^\pm(x) = |\chi_\pm\rangle\psi_n(x) = |\chi_\pm\rangle\begin{cases} \frac{1}{\sqrt{L}}\cos\left(\frac{n\pi x}{2L}\right), & n = 1,3,5,\ldots, \\ \frac{1}{\sqrt{L}}\sin\left(\frac{n\pi x}{2L}\right), & n = 2,4,6,\ldots, \end{cases} \tag{9.421}$$

where $|\chi_\pm\rangle$ represent the spinor fields corresponding to the spin-up and spin-down states, respectively:

$$|\chi_+\rangle = \left|\frac{1}{2},\frac{1}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |\chi_-\rangle = \left|\frac{1}{2},-\frac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{9.422}$$

Each energy level, $E_n = \hbar^2\pi^2n^2/(8mL^2)$, of this particle is doubly degenerate, for it corresponds to two different states.

Let us now consider the case where the particle is placed in the magnetic field $\vec{B}$. The interaction between the external magnetic field and the particle's magnetic moment $\vec{\mu}$ is given by

$$\hat{H}_p = -\vec{\mu}\cdot\vec{B} = B\mu_0\begin{cases} \sigma_z, & -L \le x \le 0, \\ \sigma_x, & 0 \le x \le L, \end{cases} \tag{9.423}$$

where we have made use of $\vec{\mu} = 2\mu_0\vec{S}/\hbar = \mu_0\vec{\sigma}$; recall that the matrices of σ_x and σ_z are

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.424}$$

To estimate the energy of this particle by means of the degenerate perturbation theory, we need to calculate first the matrix

$$\begin{pmatrix} \langle\psi_n^-|\hat{H}_p|\psi_n^-\rangle & \langle\psi_n^-|\hat{H}_p|\psi_n^+\rangle \\ \langle\psi_n^+|\hat{H}_p|\psi_n^-\rangle & \langle\psi_n^+|\hat{H}_p|\psi_n^+\rangle \end{pmatrix}, \tag{9.425}$$

where

$$\begin{aligned} \langle\psi_n^-|\hat{H}_p|\psi_n^-\rangle &= \int_{-L}^{0}|\psi_n(x)|^2\langle\chi_-|\hat{H}_p|\chi_-\rangle\,dx + \int_0^L|\psi_n(x)|^2\langle\chi_-|\hat{H}_p|\chi_-\rangle\,dx \\ &= \mu_0 B\left[\langle\chi_-|\sigma_z|\chi_-\rangle\int_{-L}^{0}|\psi_n(x)|^2\,dx + \langle\chi_-|\sigma_x|\chi_-\rangle\int_0^L|\psi_n(x)|^2\,dx\right]. \end{aligned} \tag{9.426}$$

Using $\int_{-L}^{0}|\psi_n(x)|^2\,dx = \int_0^L|\psi_n(x)|^2\,dx = \frac{1}{2}$ and since

$$\begin{aligned} \langle\chi_-|\sigma_z|\chi_-\rangle &= (0\ 1)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1, \\ \langle\chi_-|\sigma_x|\chi_-\rangle &= (0\ 1)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0, \end{aligned} \tag{9.427}$$

we have

$$\langle \psi_n^- | \hat{H}_p | \psi_n^- \rangle = -\frac{\mu_0 B}{2}. \tag{9.428}$$

Following this procedure, we can obtain the remaining matrix elements of (9.425):

$$\frac{\mu_0 B}{2} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{9.429}$$

The diagonalization of this matrix leads to

$$\left(-\frac{\mu_0 B}{2} - E^{(1)}\right)\left(\frac{\mu_0 B}{2} - E^{(1)}\right) - \left(\frac{\mu_0 B}{2}\right)^2 = 0, \tag{9.430}$$

or $E^{(1)} = \pm \mu_0 B / \sqrt{2}$. Thus, the energy of the nth excited state to first-order degenerate perturbation theory is given by

$$E_n = \frac{\hbar^2 \pi^2}{8mL^2} n^2 \pm \frac{\mu_0 B}{\sqrt{2}}. \tag{9.431}$$

The magnetic field has completely removed the degeneracy of the energy spectrum of this particle.

Problem 9.13

Consider a particle of mass m moving in the potential $V(x) = \begin{cases} +\infty, & x \leq 0, \\ \frac{1}{2} m\omega^2 x^2, & x > 0. \end{cases}$

Estimate the ground state energy of this particle using

(a) the variational method and (b) the WKB method.

Solution

(a) As seen in Problem 4.9, page 295, the ground state wave function of this potential must be selected from the harmonic oscillator wave functions that vanish at $x = 0$. Only the *odd* wave functions vanish at $x = 0$. So a trial function that, besides being zero at $x = 0$, is finite as $x \longrightarrow +\infty$ is given by

$$\psi_0(x, \alpha) = x e^{-\alpha x^2}. \tag{9.432}$$

Using the results

$$\langle \psi_0 \mid \psi_0 \rangle = \int_0^{+\infty} x^2 e^{-2\alpha x^2} \, dx = \frac{1}{8\alpha} \sqrt{\frac{\pi}{2\alpha}}, \tag{9.433}$$

$$\left\langle \psi_0 \left| \frac{1}{2} m\omega^2 x^2 \right| \psi_0 \right\rangle = \frac{1}{2} m\omega^2 \int_0^{+\infty} x^4 e^{-2\alpha x^2} \, dx = \frac{3m\omega^2}{64\alpha^2} \sqrt{\frac{\pi}{2\alpha}}, \tag{9.434}$$

$$\left\langle \psi_0 \left| -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \right| \psi_0 \right\rangle = \frac{\hbar^2}{2m} \int_0^{+\infty} \left(3\alpha x^2 - 2\alpha^2 x^4\right) e^{-2\alpha x^2} \, dx = \frac{3\hbar^2}{16m} \sqrt{\frac{\pi}{2\alpha}}, \tag{9.435}$$

we obtain the ground state energy

$$E_0(\alpha) = \frac{\langle \psi_0(\alpha) | \hat{H} | \psi_0(\alpha) \rangle}{\langle \psi_0 \mid \psi_0 \rangle} = \frac{3\hbar^2}{2m} \alpha + \frac{3m\omega^2}{8\alpha}. \tag{9.436}$$

The minimization of $E_0(\alpha)$ with respect to α yields $\alpha_0 = m\omega/(2\hbar)$ and hence $E_0(\alpha_0) = \frac{3}{2}\hbar\omega$. This energy is identical to the exact value obtained in Chapter 4.

(b) This potential contains a single rigid wall at $x = 0$. Thus the proper quantization rule for this potential is given by (9.235): $\int_0^a p\,dx = (n + \frac{3}{4})\pi\hbar$; the turning point occurs at $x = a$ with $E = \frac{1}{2}m\omega^2a^2$ and hence $a = \sqrt{2E/(m\omega^2)}$.

The calculation of $\int_0^a p\,dx$ goes as follows:

$$\int_0^a p\,dx = \int_0^a \sqrt{2mE - m^2\omega^2x^2}dx = m\omega \int_0^a \sqrt{a^2 - x^2}dx. \tag{9.437}$$

The change of variable $x = a\sin\theta$ leads to

$$\int_0^a \sqrt{a^2 - x^2}\,dx = a^2 \int_0^{\pi/2} \cos^2\theta\,d\theta = \frac{a^2}{2}\int_0^{\pi/2}(1 + \cos 2\theta)d\theta = \frac{\pi a^2}{4}; \tag{9.438}$$

hence

$$\int_0^a p\,dx = m\omega\frac{\pi a^2}{4} = \frac{\pi E}{2\omega}. \tag{9.439}$$

Since $\int_0^a p\,dx = (n + \frac{3}{4})\pi\hbar$, that is, $\pi E/(2\omega) = (n + \frac{3}{4})\pi\hbar$, we obtain

$$E_n^{WKB} = \left(2n + \frac{3}{2}\right)\hbar\omega, \qquad n = 0, 1, 2, 3, \ldots. \tag{9.440}$$

This relation is identical with the *exact* expression obtained in Chapter 4. The WKB ground state energy is thus given by $E_0^{WKB} = \frac{3}{2}\hbar\omega$.

Problem 9.14

Consider an H_2 molecule where the protons are separated by a wide distance R and both are located on the z-axis. Ignoring the spin degrees of freedom and treating the dipole–dipole interaction as a perturbation, use perturbation theory to estimate an upper limit for the ground state energy of this molecule.

Solution

Assuming the protons are fixed in space and separated by a distance R, we can write the Hamiltonian of this molecule as follows:

$$\hat{H} = \hat{H}_0 + \hat{H}_p = \hat{H}_0^A + \hat{H}_0^B + \hat{H}_p, \tag{9.441}$$

where $\hat{H}_0^A$ and $\hat{H}_0^B$ are the unperturbed Hamiltonians of atoms A and B, and $\hat{H}_p$ is

$$\hat{H}_p = \frac{e^2}{R} + \frac{e^2}{|\vec{R} + \vec{r}_A - \vec{r}_B|} - \frac{e^2}{|\vec{R} + \vec{r}_A|} - \frac{e^2}{|\vec{R} - \vec{r}_B|}, \tag{9.442}$$

where $\vec{r}_A$ and $\vec{r}_B$ are the position vectors of the electrons of atoms A and B as measured from the protons. If $R \gg a_0$, where $a_0 = \hbar^2/(\mu e^2)$ is the Bohr radius, an expansion of (9.442) in powers of $\vec{r}_A/R$ and $\vec{r}_B/R$ yields, to first nonvanishing terms, an expression of the order of $1/R^3$:

$$\hat{H}_p = \frac{e^2}{R^3}\left[\vec{r}_A \cdot \vec{r}_B - 3\frac{(\vec{r}_A \cdot \vec{R})(\vec{r}_B \cdot \vec{R})}{R^2}\right]. \tag{9.443}$$

This is the dipole–dipole interaction energy between the dipole moments of the two atoms.

Since $\vec{R} = R\hat{z}$ we can write (9.443) as

$$\hat{H}_p = \frac{e^2}{R^3}\left(\hat{X}_A\hat{X}_B + \hat{Y}_A\hat{Y}_B - 2\hat{Z}_A\hat{Z}_B\right). \tag{9.444}$$

The ground state energy and wave function of the (unperturbed) molecule are

$$E_0 = E_0^A + E_0^B = 2E_{100} = -\frac{e^2}{a_0}, \qquad |\phi_0\rangle = |\phi_0^A\rangle|\phi_0^B\rangle = |\,100\rangle_A\,|\,100\rangle_B. \tag{9.445}$$

The first-order correction to the molecule's energy, $E^{(1)} = \langle\phi_0|\hat{H}_p|\phi_0\rangle$, is given by

$$\begin{aligned} E^{(1)} = \frac{e^2}{R^3}\Big(&\langle\phi_0^A|\hat{X}_A|\phi_0^A\rangle\langle\phi_0^B|\hat{X}_B|\phi_0^B\rangle + \langle\phi_0^A|\hat{Y}_A|\phi_0^A\rangle\langle\phi_0^B|\hat{Y}_B|\phi_0^B\rangle \\ &- 2\langle\phi_0^A|\hat{Z}_A|\phi_0^A\rangle\langle\phi_0^B|\hat{Z}_B|\phi_0^B\rangle\Big). \end{aligned} \tag{9.446}$$

Since the operators $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ are odd and the states $|\phi_0^A\rangle$ and $|\phi_0^B\rangle$ are spherically symmetric, then all the terms in (9.446) are zero; hence $E^{(1)} = 0$.

Let us now calculate the second-order correction:

$$E^{(2)} = \sum_{n,l,m;n',l',m'\neq 1,0,0} \frac{\left|\langle n,l,m;n',l',m'|\hat{H}_p|\phi_0\rangle\right|^2}{2E_{100} - E_n - E_{n'}}, \tag{9.447}$$

where

$$\begin{aligned} \langle n,l,m;n',l',m'|\hat{H}_p|\phi_0\rangle = \frac{e^2}{R^3}\Big(&\langle n,l,m|X_A|1,0,0\rangle_A\langle n',l',m'|X_B|1,0,0\rangle_B \\ &+ \langle n,l,m|Y_A|1,0,0\rangle_A\langle n',l',m'|Y_B|1,0,0\rangle_B \\ &+ -2\langle n,l,m|Z_A|1,0,0\rangle_A\langle n',l',m'|Z_B|1,0,0\rangle_B\Big). \end{aligned} \tag{9.448}$$

The terms of this expression are nonzero only if $l = l' = 1$, since the $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ operators are proportional to Y_{1m}. We can evaluate $E^{(2)}$ using a crude approximation where we assume the denominator of (9.447) is constant and we take $E_n \simeq E_{n'}$. Note that, for $n \geq 2$, we have $E_{nlm} \geq E_{200}$. In this case we can rewrite (9.447) as

$$E^{(2)} \leq \frac{1}{2(E_{100} - E_{200})} \sum_{n,l,m;n',l',m'\neq 1,0,0} \left|\langle n,l,m;n',l',m'|\hat{H}_p|\phi_0\rangle\right|^2; \tag{9.449}$$

since the diagonal term is zero (i.e., $\langle 1,0,0;1,0,0|\hat{H}_p|1,0,0;1,0,0\rangle = 0$), we have

$$\begin{aligned} &\sum_{n,l,m;n',l',m'} \left|\langle n,l,m;n',l',m'|\hat{H}_p|\phi_0\rangle\right|^2 \\ &= \sum_{n,l,m;n',l',m'} \langle 1,0,0;1,0,0|\hat{H}_p|n,l,m;n',l',m'\rangle\langle n,l,m;n',l',m'|\hat{H}_p|1,0,0;1,0,0\rangle \\ &= \langle 1,0,0;1,0,0|(\hat{H}_p)^2|1,0,0;1,0,0\rangle \\ &= \frac{e^4}{R^6}\langle 1,0,0;1,0,0|\left(X_AX_B + Y_AY_B - 2Z_AZ_B\right)^2|1,0,0;1,0,0\rangle. \end{aligned} \tag{9.450}$$

The calculation of $\langle 1,0,0;1,0,0|\,(X_AX_B + Y_AY_B - 2Z_AZ_B)^2\,|1,0,0;1,0,0\rangle$ can be made easier by the use of symmetry. Due to spherical symmetry, the cross terms are zero:

$$\langle X_AY_A\rangle_A = \langle X_AZ_A\rangle_A = \langle Y_AZ_A\rangle_A = \langle X_BY_B\rangle_B = \cdots = \langle Y_BZ_B\rangle_B = 0, \tag{9.451}$$

while the others are given as follows (see (9.45)):

$$\langle X_A^2\rangle_A = \langle Y_A^2\rangle_A = \langle Z_A^2\rangle_A = \langle X_B^2\rangle_B = \langle Y_B^2\rangle_B = \langle Z_B^2\rangle_B = a_0^2, \tag{9.452}$$

where $\langle C\rangle_A = \langle\phi_0^A|C|\phi_0^A\rangle$ and $\langle D\rangle_B = \langle\phi_0^B|D|\phi_0^B\rangle$. We can thus obtain

$$\langle 1,0,0;1,0,0|\,(X_AX_B + Y_AY_B - 2Z_AZ_B)^2\,|1,0,0;1,0,0\rangle = 6a_0^4. \tag{9.453}$$

Inserting (9.453) into (9.450) and then the resultant expression into (9.449), we get

$$E^{(2)} \le \frac{\langle 1,0,0;1,0,0|(\hat{H}_p)^2|1,0,0;1,0,0\rangle}{2(E_{100}-E_{200})} = \frac{3e^4a_0^4}{R^6}\frac{1}{E_{100}-E_{200}}, \tag{9.454}$$

or

$$E^{(2)} \le -\frac{8e^2a_0^5}{R^6}, \tag{9.455}$$

because $E_{100} = -e^2/2a_0$ and $E_{200} = -e^2/8a_0$. Finally, the upper limit for the ground state energy of this molecule to second-order perturbation theory is given by

$$E_2 \le 2E_{100} - \frac{8e^2a_0^5}{R^6} \quad\Longrightarrow\quad E_2 \le -\frac{e^2}{a_0}\left(1 + 8\frac{a_0^6}{R^6}\right). \tag{9.456}$$

Problem 9.15
A proton of energy E is incident from the right on a nucleus of charge Ze. Estimate the transmission coefficient associated with the penetration of the proton inside the nucleus.

Solution
To penetrate inside the nucleus (i.e., to the left of the turning point $r = a$ as shown in Figure 9.10), the proton has to overcome the repulsive Coulomb force of the nucleus. That is, it has to tunnel through the Coulomb barrier $V(r) = Ze^2/r$. The transmission coefficient is given in the WKB approximation by (9.258), where $x_1 = a$ and $x_2 = 0$:

$$T = e^{-2\gamma}, \qquad \gamma = \frac{1}{\hbar}\int_a^0 \sqrt{2m(V(r)-E)}\,dr, \tag{9.457}$$

where a is given by $E = V(a)$: $a = Ze^2/E$. Since $V(r) = Ze^2/r$ we get

$$\gamma = \frac{1}{\hbar}\int_a^0 \sqrt{2m\left(\frac{Ze^2}{r} - E\right)}\,dr = \frac{\sqrt{2mE}}{\hbar}\int_{Ze^2/E}^0 \sqrt{\frac{Ze^2}{Er} - 1}\,dr. \tag{9.458}$$

The change of variable $x = Er/(Ze^2)$ gives

$$\gamma = \frac{Ze^2}{\hbar}\sqrt{\frac{2m}{E}}\int_0^1 \sqrt{\frac{1}{x} - 1}\,dx = \frac{Ze^2\pi}{\hbar}\sqrt{\frac{m}{2E}}; \tag{9.459}$$

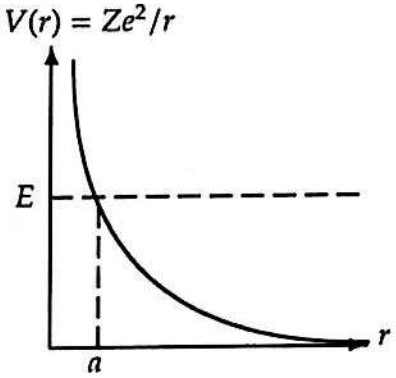


Figure 9.10: Coulomb barrier, $V(r) = Ze^2/r$, seen by a proton of energy E while approaching from the right a nucleus of charge Ze located at the origin.

in deriving this relation, we have used the integral $\int_0^1 \sqrt{1/x - 1}dx = \pi/2$.

The transmission coefficient is thus given by

$$T = e^{-2\gamma} = \exp\left\{-\frac{Ze^2\pi}{\hbar}\sqrt{\frac{2m}{E}}\right\}. \tag{9.460}$$

The value of this coefficient describes how difficult it is for a positively charged particle, such as a proton, to approach a nucleus.

Problem 9.16

Two identical particles of spin $\frac{1}{2}$ are enclosed in a one-dimensional box potential of length L with walls at $x = 0$ and $x = L$.

(a) Find the energies of the three lowest states.

(b) Then, subjecting the particles to a perturbation

$$\hat{H}_p(x_1, x_2) = -V_0 L^2 \delta\left(x_1 - \frac{L}{2}\right)\delta\left(x_2 - \frac{L}{3}\right),$$

calculate its ground state energy using first-order time-independent perturbation theory.

Solution

Since the two particles have the *same* spin, the spin wave function of the system, $\chi_s(s_1, s_2)$, must be *symmetric*, so χ_s is any one of the triplet states:

$$\chi_s = \begin{cases} |1, 1\rangle &= \left|\frac{1}{2}, \frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, \frac{1}{2}\right\rangle_2, \\ |1, 0\rangle &= \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2}, \frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_2 + \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, \frac{1}{2}\right\rangle_2\right], \\ |1, -1\rangle &= \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_2. \end{cases} \tag{9.461}$$

In addition, since this two-particle system is a system of *identical fermions*, its wave function must be *antisymmetric*. Since the spin part is symmetric, the spatial part of the wave function has to be antisymmetric:

$$\Psi(x_1, x_2) = \psi_A(x_1, x_2)\chi_s(s_1, s_2); \tag{9.462}$$

that is,

$$\psi_A(x_1,x_2) = \frac{1}{\sqrt{2}}\left[\phi_{n_1}(x_1)\phi_{n_2}(x_2) - \phi_{n_2}(x_1)\phi_{n_1}(x_2)\right]$$
$$= \frac{1}{L}\left[\sin\left(\frac{n_1\pi x_1}{L}\right)\sin\left(\frac{n_2\pi x_2}{L}\right) - \sin\left(\frac{n_2\hbar x_1}{L}\right)\sin\left(\frac{n_1\pi x_2}{L}\right)\right]. \tag{9.463}$$

The energy levels of this two-particle system are

$$E = \frac{\pi^2\hbar^2}{2mL^2}(n_1^2 + n_2^2) = E_0(n_1^2 + n_2^2), \tag{9.464}$$

where $E_0 = \pi^2\hbar^2/(2mL^2)$. Note that these energy levels are threefold degenerate because of the spin part of the wave function; that is, there are three different spin states that correspond to the same energy level $\pi^2\hbar^2(n_1^2 + n_2^2)/(2mL^2)$.

(a) Having written the general expressions for the energies and the wave functions, it is now easy to infer the energy levels and wave functions of the three lowest states. First, we should note that the ground state cannot correspond to $n_1 = n_2 = 1$, for the spatial wave function would be *zero*. The ground state corresponds then to $n_1 = 1, n_2 = 2$; its energy follows from (9.464),

$$E^{(0)} = E_0(1^2 + 2^2) = 5E_0 = \frac{5\pi^2\hbar^2}{2mL^2}, \tag{9.465}$$

and the wave function $\psi_0(x_1, x_2)$ follows from (9.463).

The first excited state corresponds to $n_1 = 1$, $n_2 = 3$. So the wave function $\psi_1(x_1, x_2)$ can be inferred from (9.463) and the energy from (9.464):

$$E^{(1)} = E_0(1^2 + 3^2) = 10E_0 = \frac{5\pi^2\hbar^2}{mL^2}. \tag{9.466}$$

The second excited state corresponds to $n_1 = 2$, $n_2 = 3$; hence the energy is given by

$$E^{(2)} = 13E_0 = \frac{13\pi^2\hbar^2}{2mL^2}. \tag{9.467}$$

(b) Introducing the perturbation $\hat{H}_p = -V_0L^2\delta(x_1 - L/2)\delta(x_2 - L/3)$, and since $\hat{H}_p$ is diagonal in the spin space, the ground state energy to first-order perturbation theory is given by

$$E = \frac{5\pi^2\hbar^2}{2mL^2} + \langle\psi_0|\hat{H}_p|\psi_0\rangle \tag{9.468}$$

where

$$\langle\psi_0|\hat{H}_p\,|\,\psi_0\rangle = \int_0^L dx_1 \int_0^L dx_2\psi_0^*(x_1,x_2)\hat{H}_p(x_1,x_2)\psi_0(x_1,x_2). \tag{9.469}$$

Since

$$\psi_0^*(x_1,x_2) = \psi_0(x_1,x_2) = \frac{1}{L}\left[\sin\left(\frac{\pi x_1}{L}\right)\sin\left(\frac{2\pi x_2}{L}\right) - \sin\left(\frac{2\pi x_1}{L}\right)\sin\left(\frac{\pi x_2}{L}\right)\right], \tag{9.470}$$

we have

$$\begin{aligned}\langle\psi_0|\hat{H}_p\,|\,\psi_0\rangle &= -\frac{V_0L^2}{L^2}\int_0^L dx_1\delta\left(x_1-\frac{L}{2}\right)\int_0^L dx_2\delta\left(x_2-\frac{L}{3}\right)\\ &\quad\times\left[\sin\left(\frac{\pi x_1}{L}\right)\sin\left(\frac{2\pi x_2}{L}\right)-\sin\left(\frac{2\pi x_1}{L}\right)\sin\left(\frac{\pi x_2}{L}\right)\right]^2\\ &= -V_0\left[\sin\left(\frac{\pi}{2}\right)\sin\left(\frac{2\pi}{3}\right)-\sin(\pi)\sin\left(\frac{\pi}{3}\right)\right]^2\\ &= -\frac{3}{4}V_0;\end{aligned} \tag{9.471}$$

hence

$$E=\frac{5\pi^2\hbar^2}{2mL^2}-\frac{3}{4}V_0. \tag{9.472}$$

Problem 9.17
Neglecting the spin–orbit interaction, find the ground state energy of a two-electron atom in these two ways:

(a) Use a first-order perturbation calculation; treat the Coulomb interaction between the two electrons as a perturbation.

(b) Use the variational method.

Compare the results and discuss the merits of the two approximation methods.

Solution
Examples of such a system are the helium atom ($Z = 2$), the singly ionized Li^+ ion ($Z = 3$), the doubly ionized Be^{2+} ion ($Z = 4$), and so on. Each electron of these systems feels the effects of two Coulomb fields: one from the Ze nucleus, $V(r) = -Ze^2/r$, and the other from the other electron, $V_{12} = e^2/r_{12} = e^2/|\vec{r}_1 - \vec{r}_2|$; here we consider the nucleus to be located at the origin and the electrons at $\vec{r}_1$ and $\vec{r}_2$. Neglecting the spin–orbit interaction, we can write the Hamiltonian of the two-electron system as

$$\hat{H}=\hat{H}_0+\hat{V}_{12}=\hat{H}_0+\frac{e^2}{|\vec{r}_1-\vec{r}_2|}, \tag{9.473}$$

where

$$\hat{H}_0=-\frac{\hbar^2}{2\mu}\left(\nabla_1^2+\nabla_2^2\right)-Ze^2\left(\frac{1}{r_1}+\frac{1}{r_2}\right) \tag{9.474}$$

is the Hamiltonian of the atom when the interaction between the two electrons is neglected.

We have seen in Chapter 8 that, when the interaction between the two electrons is neglected, the ground state energy and wave function are given by

$$E_0=-2\frac{Z^2e^2}{2a}=-27.2Z^2\,\text{eV}, \tag{9.475}$$

$$\Psi_0(\vec{r}_1,\vec{S}_1;\,\vec{r}_2,\,\vec{S}_2)=\psi_0(\vec{r}_1,\vec{r}_2)\chi_{singlet}(\vec{S}_1,\vec{S}_2), \tag{9.476}$$

where the spin part is antisymmetric,

$$\chi_{singlet}(\vec{S}_1, \vec{S}_2) = \frac{1}{\sqrt{2}}\left(\left|\frac{1}{2}, \frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_2 - \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, \frac{1}{2}\right\rangle_2\right), \tag{9.477}$$

and the spatial part is symmetric, $\psi_0(\vec{r}_1, \vec{r}_2) = \phi_{100}(\vec{r}_1)\phi_{100}(\vec{r}_2)$, with

$$\phi_{100}(\vec{r}) = R_{10}(r)Y_{00}(\Omega) = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a}\right)^{3/2} e^{-Zr_1/a}, \tag{9.478}$$

that is,

$$\psi_0(\vec{r}_1, \vec{r}_2) = \frac{1}{\pi}\left(\frac{Z}{a}\right)^3 e^{-Z(r_1+r_2)/a}. \tag{9.479}$$

(a) To calculate the ground state energy using first-order perturbation theory, we have to treat $\hat{V}_{12}$ as a perturbation. A first-order treatment yields

$$E = E_0 + \langle\psi_0|\hat{V}_{12}\,|\,\psi_0\rangle = -2\frac{Z^2e^2}{2a} + \langle\psi_0|\hat{V}_{12}\,|\,\psi_0\rangle, \tag{9.480}$$

where

$$\begin{aligned}\langle\psi_0|\hat{V}_{12}\,|\,\psi_0\rangle &= \int d^3r_1 \int d^3r_2\, \psi_0^*(\vec{r}_1, \vec{r}_2)\hat{V}_{12}\psi_0(\vec{r}_1, \vec{r}_2) \\ &= \int d^3r_1 \int d^3r_2\, |\phi_{100}(\vec{r}_1)|^2 \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} |\phi_{100}(\vec{r}_2)|^2 .\end{aligned} \tag{9.481}$$

The calculation of this integral is quite involved (I left it as an exercise); the result is

$$\langle\psi_0|\hat{V}_{12}\,|\,\psi_0\rangle = \frac{5}{8}\frac{Ze^2}{a}, \tag{9.482}$$

which, when combined with (9.480), leads to

$$E = -\frac{Ze^2}{a}\left(Z - \frac{5}{8}\right). \tag{9.483}$$

In the case of helium, $Z = 2$, we have

$$E = -108.8\,\text{eV} + 34\,\text{eV} = -74.8\,\text{eV}; \tag{9.484}$$

this result disagrees with the experimental value, $E_{exp} = -78.975\,\text{eV}$, by 4 eV or by a 5.3% relative error. Physically, this may be attributed to the fact that, in our calculation, we have not taken into account the "screening" effect: the presence of one electron tends to decrease the net charge "seen" by the other electron. Suppose electron 1 is "between" the nucleus and electron 2; then electron 2 will not "see" Z protons but $(Z-1)$ protons (i.e., electron 2 feels an effective charge $(Z-1)e$ coming from the nucleus).

(b) By analogy with the exact form of the ground state function (9.479), we can choose a trial function that takes into account the screening effect. For this, we need simply to replace Z in (9.479) by a variational parameter α:

$$\psi_0(r_1, r_2) = Ae^{-\alpha(r_1+r_2)/a}, \tag{9.485}$$

where A is a normalization. Using the integral $\int_0^\infty x^n e^{-bx}dx = n!/b^{n+1}$, we can show that $A = (\alpha/a)^3/\pi$; hence

$$\psi_\alpha(r_1, r_2) = \frac{1}{\pi}\left(\frac{\alpha}{a}\right)^3 e^{-\alpha(r_1+r_2)/a}. \tag{9.486}$$

A combination of this relation with (9.482) leads to

$$E(\alpha) = \langle\psi_\alpha|\hat{H}_0|\psi_\alpha\rangle + \langle\psi_\alpha|\hat{V}_{12}|\psi_\alpha\rangle = \langle\psi_\alpha|\hat{H}_0|\psi_\alpha\rangle + \frac{5}{8}\frac{\alpha e^2}{a}. \tag{9.487}$$

The calculation of $\langle\psi_\alpha|\hat{H}_0|\psi_\alpha\rangle$ can be simplified by writing it as

$$\begin{aligned}\langle\psi_\alpha|\hat{H}_0|\psi_\alpha\rangle &= -\frac{\hbar^2}{2\mu}\langle\psi_\alpha|\nabla_1^2 + \nabla_2^2|\psi_\alpha\rangle - Ze^2\langle\psi_\alpha|\frac{1}{r_1} + \frac{1}{r_2}|\psi_\alpha\rangle \\ &= -\frac{\hbar^2}{2\mu}\langle\psi_\alpha|\nabla_1^2 + \nabla_2^2|\psi_\alpha\rangle - \alpha e^2\langle\psi_\alpha|\frac{1}{r_1} + \frac{1}{r_2}|\psi_\alpha\rangle \\ &\quad - (Z-\alpha)e^2\langle\psi_\alpha|\frac{1}{r_1} + \frac{1}{r_2}|\psi_\alpha\rangle.\end{aligned} \tag{9.488}$$

This form is quite suggestive; since $-\hbar^2\langle\psi_0|\vec{\nabla} \mid \psi_0\rangle/(2\mu) - Ze^2\langle\psi_0|1/r \mid \psi_0\rangle = -Z^2e^2/(2a)$ we can write

$$-\frac{\hbar^2}{2\mu}\langle\psi_\alpha|\nabla_1^2 + \nabla_2^2|\psi_\alpha\rangle - \alpha e^2\langle\psi_\alpha|\frac{1}{r_1} + \frac{1}{r_2}|\psi_\alpha\rangle = -2\frac{\alpha^2e^2}{2a}. \tag{9.489}$$

Now since

$$\langle\psi_\alpha|\frac{1}{r_1}|\psi_\alpha\rangle = \langle\psi_\alpha|\frac{1}{r_2}|\psi_\alpha\rangle = 4\left(\frac{\alpha}{a}\right)^3\int_0^\infty re^{-2\alpha r/a}dr = \frac{\alpha}{a}, \tag{9.490}$$

we can reduce (9.488) to

$$\langle\psi_\alpha|\hat{H}_0|\psi_\alpha\rangle = -2\frac{\alpha^2e^2}{2a} - 2(Z-\alpha)e^2\frac{\alpha}{a}, \tag{9.491}$$

which, when combined with (9.487), leads to

$$E(\alpha) = -2\frac{\alpha^2e^2}{2a} - 2(Z-\alpha)e^2\frac{\alpha}{a} + \frac{5}{8}\frac{\alpha e^2}{a} = \left[\alpha^2 - 2\left(Z - \frac{5}{16}\right)\alpha\right]\frac{e^2}{a}. \tag{9.492}$$

The minimization of $E(\alpha)$, $dE(\alpha)/d\alpha = 0$, yields

$$\alpha_0 = Z - \frac{5}{16}; \tag{9.493}$$

hence the ground state energy is

$$E(\alpha_0) = -\left[1 - \frac{5}{8Z} + \left(\frac{5}{16Z}\right)^2\right]\frac{Z^2e^2}{a} \tag{9.494}$$

and

$$\psi(r_1, r_2) = \frac{1}{\pi}\left(\frac{Z}{a} - \frac{5}{16a}\right)^3 \exp\left[-\left(\frac{Z}{a} - \frac{5}{16a}\right)(r_1 + r_2)\right]. \tag{9.495}$$

As a numerical illustration, the ground state energy of a helium atom is obtained by substituting $Z = 2$ into (9.494). This yields $E_0 = -77.456\,\text{eV}$, in excellent agreement with the experimental value $E_{exp} = -78.975\,\text{eV}$. The variational method, which overestimates the correct result by a mere 1.9%, is significantly more accurate than first-order perturbation theory. The reason is quite obvious; while the perturbation treatment does not account for the screening effect, the variational method includes it quite accurately. The wave function (9.495) shows that the second electron does not see a charge Ze, but a lower charge $(Z-5/16)e$.

9.7 Exercises

Exercise 9.1
Calculate the energy of the nth excited state to first-order perturbation for a one-dimensional box potential of length $2L$, with walls at $x = -L$ and $x = L$, which is modified at the bottom by the following perturbations with $V_0 \ll 1$:

(a) $V_p(x) = \begin{cases} -V_0, & -L \le x \le L, \\ 0, & \text{elsewhere;} \end{cases}$

(b) $V_p(x) = \begin{cases} -V_0, & -L/2 \le x \le L/2, \\ 0, & \text{elsewhere.} \end{cases}$

Exercise 9.2
Calculate the energy of the nth excited state to first-order perturbation for a one-dimensional box potential of length $2L$, with walls at $x = -L$ and $x = L$, which is modified at the bottom by the following perturbations with $V_0 \ll 1$:

(a) $V_p(x) = \begin{cases} -V_0, & -L/2 \le x \le 0, \\ 0, & \text{elsewhere;} \end{cases}$

(b) $V_p(x) = \begin{cases} V_0, & 0 \le x \le L/2, \\ 0, & \text{elsewhere;} \end{cases}$

(c) $V_p(x) = \begin{cases} -V_0, & -L/2 \le x \le 0; \\ V_0, & 0 \le x \le L/2, \\ 0, & \text{elsewhere.} \end{cases}$

Exercise 9.3
Calculate the energy of the nth excited state to second-order perturbation and the wave function to first-order perturbation for a one-dimensional box potential of length $2L$, with walls at $x = -L$ and $x = L$, which is modified at the bottom by the following perturbations with $V_0 \ll 1$:

(a) $V_p(x) = \begin{cases} 0, & -L \le x \le 0, \\ V_0, & 0 \le x \le L; \end{cases}$ (b) $V_p(x) = \begin{cases} -V_0(1 - x^2/L^2), & |x| < L, \\ 0, & \text{elsewhere.} \end{cases}$

Exercise 9.4
Consider a system whose Hamiltonian is given by $\hat{H} = E_0 \begin{pmatrix} 3 & 2\lambda & 0 & 0 \\ 2\lambda & -3 & 0 & 0 \\ 0 & 0 & -7 & \sqrt{2}\lambda \\ 0 & 0 & \sqrt{2}\lambda & 7 \end{pmatrix}$,

where $\lambda \ll 1$.

(a) Calculate the *exact* eigenvalues of $\hat{H}$; expand each of these eigenvalues to the second power of λ.

(b) Calculate the energy eigenvalues to second-order perturbation theory and compare them with the exact results obtained in (a).

(c) Calculate the eigenstates of $\hat{H}$ up to the first-order correction.

Exercise 9.5

Consider a particle of mass m that moves in a three-dimensional potential $V(r) = kr$, where k is a constant having the dimensions of a force. Use the variational method to estimate its ground state energy; you may take $R(r) = e^{-r^2/2\alpha^2}$ as the trial radial function where α is an adjustable parameter.

Exercise 9.6

Use the WKB method to estimate the ground state energy of a particle of mass m that moves in a three-dimensional potential $V(r) = kr$, where k is a constant having the dimensions of a force.

Exercise 9.7

Consider a two-dimensional harmonic oscillator Hamiltonian:

$$\hat{H} = \frac{1}{2m}(\hat{P}_x^2 + \hat{P}_y^2) + \frac{1}{2}m\omega^2(\hat{X}^2 + \hat{Y}^2)(1 + \lambda\hat{X}\hat{Y}),$$

where $\lambda \ll 1$.

(a) Give the wave functions for the three lowest energy levels when $\lambda = 0$.

(b) Using perturbation theory, evaluate the first-order corrections of these energy levels when $\lambda \neq 0$.

Exercise 9.8

Consider a particle that has the Hamiltonian $\hat{H} = \hat{H}_0 + \lambda\hbar\omega(\hat{a}^2 + \hat{a}^{\dagger 2})$, where $\hat{H}_0$ is the Hamiltonian of a simple one-dimensional harmonic oscillator, and where $\hat{a}$ and $\hat{a}^\dagger$ are the usual annihilation and creation operators which obey $[\hat{a}, \hat{a}^\dagger] = 1$; λ is a very small real number.

(a) Calculate the ground state energy to second order in λ.

(b) Find the energy of the nth excited state, E_n, to second order in λ and the corresponding eigenstate $|\psi_n\rangle$ to first order in λ.

Exercise 9.9

Consider two identical particles of spin $\frac{1}{2}$ that are confined in an isotropic three-dimensional harmonic oscillator potential of frequency ω.

(a) Find the ground state energy and the corresponding wave function of this system when the two particles do not interact.

(b) Consider now that there exists a weakly attractive spin-dependent potential between the two particles, $V(r_1, r_2) = -kr_1r_2 - \lambda\hat{S}_{1z}\hat{S}_{2z}$, where k and λ are two small positive real numbers. Find the ground state to first-order time-independent perturbation theory.

(c) Use the variational method to estimate the ground state energy of this system of two noninteracting spin $\frac{1}{2}$ particles confined to an isotropic three-dimensional harmonic oscillator. How does your result compare with that obtained in (a).

Exercise 9.10
Two identical spin $\frac{1}{2}$ particles are confined to a one-dimensional box potential of size L with walls at $x = 0$ and $x = L$.

(a) Find the ground state energy and the first excited state energy and their respective wave functions for this system when the two particles do not interact.

(b) Consider now that there exists a *weakly* attractive potential between the two particles:

$$V_p(x) = \begin{cases} -V_0, & 0 \le x \le L/2, \\ 0, & L/2 < x \le L. \end{cases}$$

Find the ground state and first excited state energies to first-order perturbation theory.

(c) Find numerical values for the ground state and first excited state energies calculated in (a) in the case where $L = 10^{-10}$ m, $V_0 = 2$ eV, and the mass of each individual particle is to be taken equal to the electron mass. Compare the sizes of the first order energy corrections with the ground state energy and the first excited state energy (you may simply calculate the ratios between the first-order corrections with the ground state and first excited state energies).

Exercise 9.11
Consider an isotropic three-dimensional harmonic oscillator.

(a) Find the energy of the first excited state and the different states corresponding to this energy.

(b) If we now subject this oscillator to a perturbation $\hat{V}_p(x, y) = -\lambda \hat{x}\hat{y}$, where λ is a small real number, find the energy of the first excited state to first-order degenerate time-independent perturbation theory. *Hint:* You may use $\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}_x + \hat{a}_x^\dagger)$, $\hat{y} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}_y + \hat{a}_y^\dagger)$, $\hat{z} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}_z + \hat{a}_z^\dagger)$, with $\hat{a}_x|n_x\rangle = \sqrt{n_x}|n_x - 1\rangle$, and $\hat{a}_x^\dagger|n_x\rangle = \sqrt{n_x + 1}|n_x + 1\rangle$.

Exercise 9.12
Use the variational method to estimate the ground state energy of the spherical harmonic oscillator by means of the following radial trial functions:

(a) $R(r) = Ae^{-ar^2}$ and

(b) $R(r) = Ae^{-ar}$, where A is a normalization constant that needs to be found in each case and a is an adjustable parameter.

(c) Using the fact that the exact ground state energy is $E_0^{exact} = 3\hbar\omega/2$, find the relative errors corresponding to the energies derived in (a) and (b).

Exercise 9.13
Consider a particle of mass m that is bouncing vertically and elastically on a smooth reflecting floor in the Earth's gravitational field

$$V(z) = \begin{cases} mgz, & z > 0, \\ +\infty, & z \le 0, \end{cases}$$

where g is a constant (the acceleration due to gravity). Use the variational method to estimate the ground state energy of this particle by means of the trial wave function, $\psi(z) = z\exp(-az^4)$, where a is an adjustable parameter that needs to be determined. Compare your result with the exact value $E_0^{exact} = 2.338\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3}$ by calculating the relative error.

Exercise 9.14
Calculate the energy of the ground state to first-order perturbation for a particle which is moving in a one-dimensional box potential of length L, with walls at $x = 0$ and $x = L$, when a weak potential $\hat{H}_p = \lambda x^2$ is added, where $\lambda \ll 1$.

Exercise 9.15
Consider a semiclassical system whose energy is given by

$$E = a^2 + \frac{1}{2}\left(\frac{b^2}{4} - a^2\right)p^2 + \frac{1}{2}\left(\frac{4a^2}{b^2/4 - a^2}\right)q^2,$$

where a and b are positive, real constants. Use the Bohr–Sommerfeld quantization rule to extract the expression of the bound state energy E_n for the nth excited state in terms of a.

Exercise 9.16
Use the variational method to estimate the ground state energy of a particle of mass m that is moving in a one-dimensional potential $V(x) = V_0x^4$; you may use the trial function $\psi_0(x, \alpha) = Ae^{-\alpha x^2/2}$, where A is the normalization constant and α is an adjustable parameter that needs to be determined.

Exercise 9.17
Consider a particle of mass m which is moving in a one-dimensional potential $V(x) = V_0x^4$. Estimate the ground state energy of this particle by means of the WKB method.

Exercise 9.18
Using $\phi_{100}(\vec{r}) = (1/\sqrt{\pi})(Z/a)^{3/2}e^{-Zr/a}$, show that

$$\int \phi_{100}^2(\vec{r}_1)\frac{1}{|\vec{r}_1 - \vec{r}_2|}\phi_{100}^2(\vec{r}_2)d^3r_1d^3r_2 = \frac{5Z}{8a}.$$

Exercise 9.19
Calculate the ground state energy of the doubly ionized beryllium atom Be^{2+} by means of the following two methods and then compare the two results:

(a) a first-order perturbation theory treatment,

(b) the variational method with a trial function $\psi(r_1, r_2) = A\exp[-\alpha(r_1 + r_2)/a]$, where A is the normalization constant, α is an adjusted parameter, and a is the Bohr radius.

Exercise 9.20
Use the variational method to estimate the energy of the second excited state of a particle of mass m moving in a one-dimensional infinite well with walls at $x = 0$ and $x = L$. Calculate the relative error between your result and the exact value (recall that the energy of the second excited state is given by $E_3^{exact} = 9\pi^2\hbar^2/(2mL^2)$).

Exercise 9.21
Consider a spinless particle of orbital angular momentum $l = 1$ whose Hamiltonian is

$$\hat{H}_0 = \frac{\mathcal{E}}{\hbar^2}(\hat{L}_x^2 - \hat{L}_y^2),$$

where $\mathcal{E}$ is a constant having the dimensions of energy.

(a) Calculate the exact energy levels and the corresponding eigenstates of this particle.

(b) We now add a perturbation $\hat{H}_p = \alpha \hat{L}_z/\hbar$, where α is a small constant (small compared to $\mathcal{E}$) having the dimensions of energy. Calculate the energy levels of this particle to second-order perturbation theory.

(c) Diagonalize the matrix of $\hat{H} = \hat{H}_0 + \hat{H}_p$ and find the exact energy eigenvalues. Then expand each eigenvalue to second power in α and compare them with the results derived from perturbation theory in (b).

Exercise 9.22

Consider a system whose Hamiltonian is given by $\hat{H} = E_0 \begin{pmatrix} -5 & 3\lambda & 0 & 0 \\ 3\lambda & 5 & 0 & 0 \\ 0 & 0 & 8 & -\lambda \\ 0 & 0 & -\lambda & -8 \end{pmatrix}$, where $\lambda \ll 1$.

(a) By decomposing this Hamiltonian into $\hat{H} = \hat{H}_0 + \hat{H}_p$, find the eigenvalues and eigenstates of the unperturbed Hamiltonian $\hat{H}_0$.

(b) Diagonalize $\hat{H}$ to find the exact eigenvalues of $\hat{H}$; expand each eigenvalue to the second power of λ.

(c) Using first- and second-order nondegenerate perturbation theory, find the approximate eigenergies of $\hat{H}$. Compare these with the exact values obtained in (b).

Exercise 9.23

Estimate the ground state energy of the hydrogen atom by means of the variational method using the following two trial functions, find the relative errors, compare the two results, and discuss the merit of each trial function.

(a) $\phi_\alpha(r) = \begin{cases} 1 - r/\alpha, & r \le \alpha, \\ 0, & r > \alpha, \end{cases}$

where α is an adjustable parameter. Find a relation between α_{min} and the Bohr radius.

(b) $\phi_\alpha(r) = Ae^{-\alpha r^2}$.

Exercise 9.24

(a) Calculate to first-order perturbation theory the energy of the nth excited state of a one-dimensional harmonic oscillator which is subject to the following small perturbation: $\hat{H}_p = \lambda(V_0\hat{x}^3 + V_1\hat{x}^4)$, where V_0 and V_1 are constants and $\lambda \ll 1$.

(b) Use the relation derived in (a) to find the energies of the three lowest states (i.e., $n = 0, 1, 2$) to first-order perturbation theory.

Exercise 9.25

Use the trial function $\psi_0(x, \alpha) = \begin{cases} A(\alpha^2 - x^2)^2, & |x| \le \alpha, \\ 0, & |x| \ge \alpha, \end{cases}$

to estimate the ground state energy of a one-dimensional harmonic oscillator by means of the variational method; α is an adjustable parameter and A is the normalization constant. Calculate the relative error and assess the accuracy of the result.

Exercise 9.26

Use the WKB approximation to estimate the transmission coefficient of a particle of mass m

and energy E moving in the following potential barrier:

$$V(x) = \begin{cases} V_0(x/a + 1), & -a < x < 0, \\ V_0(1 - x/a), & 0 < x < a, \\ 0, & \text{elsewhere,} \end{cases}$$

with $0 < E < V_0$; sketch this potential.

Exercise 9.27
Use the variational method to estimate the energy of the ground state of a one-dimensional harmonic oscillator taking the following trial function:

$$\psi_0(x, \alpha) = A\,(1 + \alpha|x|)\,e^{-\alpha|x|},$$

where α is an adjustable parameter and A is the normalization constant.

Exercise 9.28
Use the WKB approximation to estimate the transmission coefficient of a particle of mass m and energy E moving in the following potential barrier:

$$V(x) = \begin{cases} V_0(1 - x^2/a^2) & |x| < a \\ 0 & |x| > a, \end{cases}$$

where $0 < E \leq V_0$.

Exercise 9.29
Use the WKB approximation to find the energy levels of a particle of mass m moving in the following potential:

$$V(x) = \begin{cases} V_0(x^2/a^2 - 1), & |x| < a, \\ 0, & |x| > a. \end{cases}$$

Exercise 9.30
A particle of mass m is moving in a one-dimensional harmonic oscillator potential, $V(x) = m\omega^2x^2/2$. Calculate

(a) the ground state energy, and

(b) the first excited state energy

to first-order perturbation theory when a small perturbation $\hat{H}_p = \lambda\delta(x)$ is added to the potential, with $\lambda \ll 1$.

Exercise 9.31
A particle of mass m is moving in a one-dimensional harmonic oscillator potential, $V(x) = m\omega^2x^2/2$. Calculate

(a) the ground state energy and

(b) the first excited state energy

to first-order perturbation theory when a small perturbation $\hat{H}_p = \lambda x^6$ is added to the potential, with $\lambda \ll 1$.

Exercise 9.32
A particle of mass m is moving in a three-dimensional harmonic oscillator potential, $V(x) = m\omega^2(x^2 + y^2 + z^2)/2$. Calculate the energy of the nth excited state to first-order perturbation theory when a small perturbation $\hat{H}_p = \lambda\hat{X}^2\hat{Y}^4\hat{Z}^2$ is added to the potential, with $\lambda \ll 1$.

Exercise 9.33
Use the following two trial functions:

$$\text{(a) } Ae^{-\alpha|x|}, \qquad \text{(b) } A(1+\alpha|x|)e^{-\alpha|x|},$$

to estimate, by means of the variational method, the ground state energy of a particle of mass m moving in a one-dimensional potential $V(x) = \lambda|x|$; α is a scale parameter, λ is a constant, and A is the normalization constant. Compare the results obtained.

Exercise 9.34
Three distinguishable particles of equal mass m are enclosed in a one-dimensional box potential with rigid walls at $x = 0$ and $x = L$. If the three particles are subject to a weak, short-range attractive potential

$$\hat{H}_p = -V_0\left[\delta(x_1 - x_2) + \delta(x_2 - x_3) + \delta(x_3 - x_1)\right],$$

use first-order perturbation theory to calculate the system's energy levels of
(a) the ground state, and
(b) the first excited state.

Exercise 9.35
Three distinguishable particles of equal mass m are in a one-dimensional harmonic oscillator potential $\hat{H}_0 = \sum_{i=1}^{3}(p_i^2/2m + \frac{1}{2}m\omega^2 x_i^2)$. If the three particles are subject to a weak, short-range attractive potential

$$\hat{H}_p = -V_0\left[\delta(x_1 - x_2) + \delta(x_2 - x_3) + \delta(x_3 - x_1)\right],$$

use first-order perturbation theory to calculate the system's energy levels of
(a) the ground state and
(b) the first-excited state.

Exercise 9.36
Consider a positronium which is subject to a *weak* static magnetic field in the xz-plane, $\vec{B} = B(\hat{\imath} + \hat{k})$, where B is a small constant. Neglecting the spin–orbit interaction, calculate the energy levels of the $n = 2$ states to first-order perturbation.

Exercise 9.37
Consider a spherically symmetric top with principal moments of inertia I.
(a) Find the energy levels of the top.
(b) Assuming that the top is in the $l = 1$ angular momentum state, find its energy to first-order perturbation theory when a weak perturbation, $\hat{H}_p = \frac{\mathcal{E}}{I}(\hat{L}_x^2 - \hat{L}_y^2)$, is added where $\mathcal{E} \ll 1$.

Exercise 9.38
Estimate the approximate values of the ground state energy of a particle of mass m moving in the potential $V(x) = V_0|x|$, where $V_0 > 0$, by means of: (a) the variational method and (b) the WKB approximation. Compare the two results.

Exercise 9.39
Calculate to first-order perturbation theory the relativistic correction to the ground state of a spinless particle of mass m moving in a one-dimensional harmonic oscillator potential.

Hint: You need first to show that the Hamiltonian can be written as $\hat{H} = \hat{H}_0 + \hat{H}_p$, where $\hat{H}_0 = \hat{P}^2/(2m) + m\omega^2\hat{X}^2/2$ and $\hat{H}_p = -\hat{P}^4/(8m^3c^2)$ is the leading relativistic correction term which can be treated as a perturbation.

Exercise 9.40
Consider a hydrogen atom which is subject to a small perturbation $\hat{H}_p = \lambda r^2$. Use a first-order perturbation theory to calculate the energy corrections to
(a) the ground state and
(b) the 2p state.

Exercise 9.41
(a) Calculate to first-order perturbation theory the contribution due to the spin–orbit interaction for the nth excited state for a positronium atom.
(b) Use the result of part (a) to obtain numerical values for the spin–orbit correction terms for the 2p level and compare them to the energy of $n = 2$.

Exercise 9.42
Ignoring the spin of the electron, calculate to first-order perturbation theory the energy of the $n = 2$ level of a hydrogen atom when subject to a *weak* quadrupole field $\hat{H}_p = iQ(y^2 - x^2)$, where Q is a small, real number $Q \ll 1$.

Exercise 9.43
Calculate the energy levels of the $n = 2$ states of positronium in a weak external electrical field $\vec{\mathcal{E}}$ along the z-axis: $\vec{\mathcal{E}} = \mathcal{E}\hat{k}$; positronium consists of an electron and a positron bound by the electric interaction.

Exercise 9.44
(a) Calculate to first-order perturbation theory the contributions due to the spin–orbit interaction for a hydrogen-like ion having Z protons.
(b) Use the result of part (a) to find the spin-orbit correction for the 2p state of a C^{5+} carbon ion and compare it with the energy of the $n = 2$ level.

Exercise 9.45
Two identical particles of spin $\frac{1}{2}$ are enclosed in a cubical box of side L.
(a) Calculate to first-order perturbation theory the ground state energy when the two particles are subject to *weak* short-range, attractive interaction:

$$\hat{V}(\vec{r}_1 - \vec{r}_2) = -\frac{4}{3}\pi a^3 V_0 \delta(\vec{r}_1 - \vec{r}_2).$$

(b) Find a numerical value for the energy derived in (a) for $L = 10^{-10}$ m, $a = 10^{-12}$ m, $V_0 = 10^{-3}$ eV, and the mass of each individual particle is to be taken to be the mass of the electron.

Chapter 10

Time-Dependent Perturbation Theory

10.1 Introduction

We have dealt so far with Hamiltonians that do not depend explicitly on time. In nature, however, most quantum phenomena are governed by time-dependent Hamiltonians. In this chapter we are going to consider approximation methods treating Hamiltonians that depend explicitly on time.

To study the structure of molecular and atomic systems, we need to know how electromagnetic radiation interacts with these systems. Molecular and atomic spectroscopy deals in essence with the absorption and emission of electromagnetic radiation by molecules and atoms. As a system absorbs or emits radiation, it undergoes transitions from one state to another.

Time-dependent perturbation theory is most useful for studying processes of absorption and emission of radiation by atoms or, more generally, for treating the transitions of quantum systems from one energy level to another.

10.2 The Pictures of Quantum Mechanics

As seen in Chapter 2, there are many representations of wave functions and operators in quantum mechanics. The connection between the various representations is provided by unitary transformations. Each class of representation, also called a *picture*, differs from others in the way it treats the time evolution of the system.

In this section we look at the pictures encountered most frequently in quantum mechanics: the Schrödinger picture, the Heisenberg picture, and the interaction picture. The Schrödinger picture is useful when describing phenomena with time-independent Hamiltonians, whereas the interaction and Heisenberg pictures are useful when describing phenomena with time-dependent Hamiltonians.

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

10.2.1 The Schrödinger Picture

In describing quantum dynamics, we have been using so far the Schrödinger picture in which *state vectors depend explicitly on time, but operators do not*:

$$\boxed{i\hbar\frac{d}{dt}\,|\,\psi(t)\rangle = \hat{H}\,|\,\psi(t)\rangle,} \tag{10.1}$$

where $|\,\psi(t)\rangle$ denotes the state of the system in the Schrödinger picture. We have seen in Chapter 3 that the time evolution of a state $|\,\psi(t)\rangle$ can be expressed by means of the propagator, or time-evolution operator, $\hat{U}(t,t_0)$, as follows:

$$\boxed{|\,\psi(t)\rangle = \hat{U}(t,t_0)\,|\,\psi(t_0)\rangle,} \tag{10.2}$$

with

$$\hat{U}(t,t_0) = e^{-i(t-t_0)\hat{H}/\hbar}. \tag{10.3}$$

The operator $\hat{U}(t,t_0)$ is unitary,

$$\hat{U}^{\dagger}(t,t_0)\hat{U}(t,t_0) = I, \tag{10.4}$$

and satisfies these properties:

$$\hat{U}(t,t) = I, \tag{10.5}$$

$$\hat{U}^{\dagger}(t,t_0) = \hat{U}^{-1}(t,t_0) = \hat{U}(t_0,t), \tag{10.6}$$

$$\hat{U}(t_1,t_2)\hat{U}(t_2,t_3) = \hat{U}(t_1,t_3). \tag{10.7}$$

10.2.2 The Heisenberg Picture

In this picture the time dependence of the state vectors is completely frozen. The Heisenberg picture is obtained from the Schrödinger picture by applying $\hat{U}$ on $|\,\psi(t)\rangle_H$:

$$|\,\psi(t)\rangle_H = \hat{U}^{\dagger}(t)\,|\,\psi(t)\rangle = |\,\psi(0)\rangle, \tag{10.8}$$

where $|\,\psi(t)\rangle$ and $\hat{U}^{\dagger}(t)$ can be obtained from (10.2) and (10.3), respectively, by setting $t_0 = 0$: $\hat{U}^{\dagger}(t) = \hat{U}^{\dagger}(t,t_0=0) = e^{it\hat{H}/\hbar}$ and $|\,\psi(t)\rangle = \hat{U}(t)\,|\,\psi(0)\rangle$, with $\hat{U}(t) = e^{-it\hat{H}/\hbar}$. Thus, we can rewrite (10.8) as

$$\boxed{|\,\psi(t)\rangle_H = e^{it\hat{H}/\hbar}\,|\,\psi(t)\rangle.} \tag{10.9}$$

As $|\,\psi\rangle_H$ is frozen in time we have: $d\,|\,\psi\rangle_H/dt = 0$. Let us see how the expectation value of an operator $\hat{A}$ in the state $|\,\psi(t)\rangle$ evolves in time:

$$\langle\psi(t)|\hat{A}|\psi(t)\rangle = \langle\psi(0)|e^{it\hat{H}/\hbar}\hat{A}e^{-it\hat{H}/\hbar}|\psi(0)\rangle = \langle\psi(0)|\hat{A}_H(t)|\psi(0)\rangle = {}_H\langle\psi|\hat{A}_H(t)|\psi\rangle_H, \tag{10.10}$$

where $\hat{A}_H(t)$ is given by

$$\boxed{\hat{A}_H(t) = \hat{U}^{\dagger}(t)\hat{A}\hat{U}(t) = e^{it\hat{H}/\hbar}\hat{A}e^{-it\hat{H}/\hbar}.} \tag{10.11}$$

Remark: An idea that helps in clarifying the difference between the Schrödinger and the Heisenberg pictures is this: the expectation value of an operator is the same in both pictures. From (10.10) and (10.11) we see that both the Schrödinger and the Heisenberg pictures coincide at $t = 0$, since $|\psi(0)\rangle_H = |\psi(0)\rangle$ and $\hat{A}_H(0) = \hat{A}$. That is, and as shown in (10.10), the expectation value of the (time-independent) operator $\hat{A}$ in the (time-dependent) sate $|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle$ is equal to the expectation value of the (time-dependent) operator $e^{it\hat{H}/\hbar}\hat{A}e^{-it\hat{H}/\hbar} = \hat{U}^\dagger(t)\hat{A}\hat{U}(t)$ in the (time-independent) state $|\psi(0)\rangle$:

$$\langle\psi(t)|\hat{A}|\psi(t)\rangle = \langle\psi(0)|\hat{A}_H(t)|\psi(0)\rangle, \tag{10.12}$$

which can be rewritten in the following more suggestive from

$$\left(\langle\psi(0)|\hat{U}^\dagger(t)\right)\hat{A}\left(\hat{U}(t)|\psi(0)\rangle\right) = \langle\psi(0)|\left(\hat{U}^\dagger(t)\,\hat{A}\,\hat{U}(t)\right)|\psi(0)\rangle. \tag{10.13}$$

10.2.2.1 The Heisenberg Equation of Motion

Let us now derive the equation of motion that regulates the time evolution of operators within the Heisenberg picture. Assuming that $\hat{A}$ does not depend explicitly on time (i.e., $\partial\hat{A}/\partial t = 0$) and since $\hat{U}(t)$ is unitary, we have

$$\begin{aligned}\frac{d\hat{A}_H(t)}{dt} &= \frac{\partial\hat{U}^\dagger(t)}{dt}\hat{A}\hat{U}_{(t)} + \hat{U}^\dagger(t)\hat{A}\frac{\partial\hat{U}(t)}{\partial t} = -\frac{1}{i\hbar}\hat{U}^\dagger\hat{H}\hat{U}\hat{U}^\dagger\hat{A}\hat{U} + \frac{1}{i\hbar}\hat{U}^\dagger\hat{A}\hat{U}\hat{U}^\dagger\hat{H}\hat{U} \\ &= \frac{1}{i\hbar}\left[\hat{A}_H, \hat{U}^\dagger\hat{H}\hat{U}\right],\end{aligned} \tag{10.14}$$

where we have used (10.3) to write $\partial\hat{U}(t)/\partial t = \hat{H}\hat{U}/i\hbar$ and $\partial\hat{U}^\dagger(t)/\partial t = -\hat{U}^\dagger\hat{H}/i\hbar$. Since $\hat{U}(t)$ and $\hat{H}$ commute, we have $\hat{U}^\dagger(t)\hat{H}\hat{U}(t) = H$; hence we can rewrite (10.14) as

$$\boxed{\frac{d\hat{A}_H}{dt} = \frac{1}{i\hbar}\left[\hat{A}_H, \hat{H}\right].} \tag{10.15}$$

This is the *Heisenberg equation of motion*. It plays the role of the Schrödinger equation within the Heisenberg picture. Since the Schrödinger and Heisenberg pictures are equivalent, we can use either picture to describe the quantum system under consideration. The Heisenberg equation (10.15), however, is in general difficult to solve.

Note that the structure of the Heisenberg equation (10.15) is similar to the classical equation of motion of a variable A that does not depend explicitly on time $dA/dt = \{A, H\}$, where $\{A, H\}$ is the Poisson bracket between A and H (see Chapter 3).

10.2.3 The Interaction Picture

The interaction picture, also called the *Dirac picture*, is useful to describe quantum phenomena with Hamiltonians that depend explicitly on time. In this picture *both state vectors and operators evolve in time*. We need, therefore, to find the equation of motion for the state vectors and for the operators.

10.2.3.1 Equation of Motion for the State Vectors

State vectors in the interaction picture are defined in terms of the Schrödinger states $|\psi(t)\rangle$ by

$$\boxed{|\psi(t)\rangle_I = e^{it\hat{H}_0/\hbar}\,|\psi(t)\rangle.} \tag{10.16}$$

If $t = 0$ we have $|\psi(0)\rangle_I = |\psi(0)\rangle$. The time evolution of $|\psi(t)\rangle$ is governed by the Schrödinger equation (10.1) with $\hat{H} = \hat{H}_0 + \hat{V}$ where $\hat{H}_0$ is time independent, but $\hat{V}$ may depend on time.

To find the time evolution of $|\psi(t)\rangle_I$, we need the time derivative of (10.16):

$$\begin{aligned} i\hbar\frac{d\,|\psi(t)\rangle_I}{dt} &= -\hat{H}_0 e^{it\hat{H}_0/\hbar}\,|\psi(t)\rangle + e^{it\hat{H}_0/\hbar}\left(i\hbar\frac{d\,|\psi(t)\rangle}{dt}\right) \\ &= -\hat{H}_0\,|\psi(t)\rangle_I + e^{it\hat{H}_0/\hbar}\hat{H}\,|\psi(t)\rangle, \end{aligned} \tag{10.17}$$

where we have used (10.1). Since $\hat{H} = \hat{H}_0 + \hat{V}$ and

$$e^{iH_0t/\hbar}\hat{V} = \left(e^{it\hat{H}_0/\hbar}\hat{V}e^{-it\hat{H}_0/\hbar}\right)e^{it\hat{H}_0/\hbar} = \hat{V}_I(t)e^{it\hat{H}_0/\hbar}, \tag{10.18}$$

with

$$\hat{V}_I(t) = e^{it\hat{H}_0/\hbar}\hat{V}e^{-it\hat{H}_0/\hbar}, \tag{10.19}$$

we can rewrite (10.17) as

$$i\hbar\frac{d\,|\psi(t)\rangle_I}{dt} = -\hat{H}_0\,|\psi(t)\rangle_I + \hat{H}_0 e^{it\hat{H}_0/\hbar}\,|\psi(t)\rangle + \hat{V}_I(t)e^{it\hat{H}_0/\hbar}\,|\psi(t)\rangle, \tag{10.20}$$

or

$$\boxed{i\hbar\frac{d\,|\psi(t)\rangle_I}{dt} = \hat{V}_I(t)\,|\psi(t)\rangle_I.} \tag{10.21}$$

This is the Schrödinger equation in the interaction picture. It shows that the time evolution of the state vector is governed by the interaction $\hat{V}_I(t)$.

10.2.3.2 Equation of Motion for the Operators

The interaction representation of an operator $\hat{A}_I(t)$ is given, as shown in (10.19), in terms of its Schrödinger representation by

$$\boxed{\hat{A}_I(t) = e^{i\hat{H}_0t/\hbar}\hat{A}e^{-i\hat{H}_0t/\hbar}.} \tag{10.22}$$

Calculating the time derivative of $\hat{A}_I(t)$ and since $\partial\hat{A}/\partial t = 0$, we can show the time evolution of $\hat{A}_I(t)$ is governed by $\hat{H}_0$:

$$\boxed{\frac{d\hat{A}_I(t)}{dt} = \frac{1}{i\hbar}\left[\hat{A}_I(t),\,\hat{H}_0\right].} \tag{10.23}$$

This equation is similar to the Heisenberg equation of motion (10.15), except that $\hat{H}$ is replaced by $\hat{H}_0$. The basic difference between the Heisenberg and interaction pictures can be inferred from a comparison of (10.9) with (10.16), and (10.11) with (10.22): in the Heisenberg picture it is $\hat{H}$ that appears in the exponents, whereas in the interaction picture it is $\hat{H}_0$ that appears.

In conclusion, we have seen that, within the Schrödinger picture, the states depend on time but not the operators; in the Heisenberg picture, only operators depend explicitly on time, state vectors are frozen in time. The interaction picture, however, is intermediate between the Schrödinger and the Heisenberg pictures, since both state vectors and operators evolve with time.

10.3 Time-Dependent Perturbation Theory

We consider here only those phenomena that are described by Hamiltonians which can be split into two parts, a time-independent part $\hat{H}_0$ and a time-dependent part $\hat{V}(t)$ that is small compared to $\hat{H}_0$:

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t), \tag{10.24}$$

where $\hat{H}_0$, which describes the system when unperturbed, is assumed to have exact solutions that are known. Such splitting of the Hamiltonian is encountered in the following typical problem. Consider a system which, when unperturbed, is described by a time-independent Hamiltonian $\hat{H}_0$ whose solutions—the eigenvalues E_n and eigenstates $|\psi_n\rangle$—are known,

$$\hat{H}_0 |\psi_n\rangle = E_n |\psi_n\rangle, \tag{10.25}$$

and whose most general state vectors are given by stationary states

$$|\Psi_n(t)\rangle = e^{-it\hat{H}_0/\hbar} |\psi_n\rangle = e^{-iE_nt/\hbar} |\psi_n\rangle. \tag{10.26}$$

In the time interval $0 \le t \le \tau$ we subject the system to an external time-dependent perturbation, $\hat{V}(t)$, that is small compared to $\hat{H}_0$:

$$\hat{V}(t) = \begin{cases} \hat{V}(t), & 0 \le t \le \tau, \\ 0, & t < 0, \quad t > \tau. \end{cases} \tag{10.27}$$

During the time interval $0 \le t \le \tau$, the Hamiltonian of the system is $\hat{H} = \hat{H}_0 + \hat{V}(t)$ and the corresponding Schrödinger equation is

$$i\hbar \frac{d|\Psi(t)\rangle}{dt} = (\hat{H}_0 + \hat{V}(t))|\Psi(t)\rangle, \tag{10.28}$$

where $\hat{V}(t)$ characterizes the interaction of the system with the external source of perturbation.

How does $\hat{V}(t)$ affect the system? When the system interacts with $\hat{V}(t)$, it either absorbs or emits energy. This process inevitably causes the system to undergo transitions from one unperturbed eigenstate to another. The main task of time-dependent perturbation theory consists of answering this question: If the system is initially in an (unperturbed) eigenstate $|\psi_i\rangle$ of $\hat{H}_0$, what is the probability that the system will be found at a later time in another unperturbed eigenstate $|\psi_f\rangle$?

To prepare the ground for answering this question, we need to look for the solutions of the Schrödinger equation (10.28). The standard method to solve (10.28) is to expand $|\Psi(t)\rangle$ in terms of an expansion coefficient $c_n(t)$:

$$|\Psi(t)\rangle = \sum_n c_n(t) e^{-iE_nt} |\psi_n\rangle, \tag{10.29}$$

and then insert this into (10.28) to find $c_n(t)$ to various orders in the approximation. Instead of following this procedure, and since we are dealing with time-dependent potentials, it is more convenient to solve (10.28) in the interaction picture (10.21):

$$i\hbar\frac{d\,|\,\Psi(t)\rangle_I}{dt} = \hat{V}_I(t)\,|\,\Psi(t)\rangle_I, \tag{10.30}$$

where $|\,\Psi(t)\rangle_I = e^{it\hat{H}_0/\hbar}\,|\,\Psi(t)\rangle$ and $\hat{V}_I(t) = e^{it\hat{H}_0/\hbar}\hat{V}(t)e^{-it\hat{H}_0/\hbar}$. The time evolution equation $|\,\Psi(t)\rangle = \hat{U}(t,t_i)\,|\,\Psi(t_i)\rangle$ may be written in the interaction picture as

$$|\,\Psi(t)\rangle_I = e^{it\hat{H}_0/\hbar}\,|\,\Psi(t)\rangle = e^{it\hat{H}_0/\hbar}\hat{U}(t,t_i)\,|\,\Psi(t_i)\rangle = e^{it\hat{H}_0/\hbar}\hat{U}(t,t_i)e^{-it\hat{H}_0t_i/\hbar}\,|\,\Psi(t_i)\rangle_I, \tag{10.31}$$

or as

$$|\,\Psi(t)\rangle_I = \hat{U}_I(t,t_i)\,|\,\Psi(t_i)\rangle_I, \tag{10.32}$$

where the time evolution operator is given in the interaction picture by

$$\hat{U}_I(t,t_i) = e^{it\hat{H}_0/\hbar}\hat{U}(t,t_i)e^{-it\hat{H}_0t_i/\hbar}. \tag{10.33}$$

Inserting (10.32) into (10.30) we end up with

$$i\hbar\frac{d\hat{U}_I(t,t_i)}{dt} = \hat{V}_I(t)\hat{U}_I(t,t_i). \tag{10.34}$$

The solutions of this equation, with the initial condition $\hat{U}_I(t_i,t_i) = \hat{I}$, are given by the *integral equation*

$$\hat{U}_I(t,t_i) = 1 - \frac{i}{\hbar}\int_{t_i}^{t}\hat{V}_I(t')\hat{U}_I(t',t_i)\,dt'. \tag{10.35}$$

Time-dependent perturbation theory provides *approximate* solutions to this integral equation. This consists in assuming that $\hat{V}_I(t)$ is *small and then proceeding iteratively*. The first-order approximation is obtained by inserting $\hat{U}_I(t',t_i) = 1$ in the integral sign of (10.35), leading to $\hat{U}_I^{(1)}(t,t_i) = 1 - (i/\hbar)\int_{t_i}^{t}\hat{V}_I(t')\,dt'$. Substituting $\hat{U}_I(t',t_i) = \hat{U}_I^{(1)}(t',t_i)$ in the integral sign of (10.35) we get the second-order approximation:

$$\hat{U}_I^{(2)}(t,t_i) = 1 - \frac{i}{\hbar}\int_{t_i}^{t}\hat{V}_I(t')\,dt' + \left(-\frac{i}{\hbar}\right)^2\int_{t_i}^{t}\hat{V}_I(t_1)\,dt_1\int_{t_i}^{t_1}\hat{V}_I(t_2)\,dt_2. \tag{10.36}$$

The third-order approximation is obtained by substituting $\hat{U}_I^{(2)}(t,t_i)$ into (10.35), and so on. A repetition of this iterative process yields

$$\begin{aligned}\hat{U}_I(t,t_i) = 1 &- \frac{i}{\hbar}\int_{t_i}^{t}\hat{V}_I(t')dt' + \left(-\frac{i}{\hbar}\right)^2\int_{t_i}^{t}\hat{V}_I(t_1)dt_1\int_{t_i}^{t_1}\hat{V}_I(t_2)dt_2 + \cdots \\ &+ \left(-\frac{i}{\hbar}\right)^n\int_{t_i}^{t}\hat{V}_I(t_1)dt_1\int_{t_i}^{t_1}\hat{V}_I(t_2)dt_2\int_{t_i}^{t_2}\hat{V}_I(t_3)dt_3\cdots\int_{t_i}^{t_{n-1}}\hat{V}_I(t_n)dt_n + \cdots.\end{aligned} \tag{10.37}$$

This series, known as the *Dyson series*, allows for the calculation of the state vector up to the desired order in the perturbation.

We are now equipped to calculate the transition probability. It may be obtained by taking the matrix elements of (10.37) between the eigenstates of $\hat{H}_0$. Time-dependent perturbation theory, where one assumes knowledge of the solutions of the unperturbed eigenvalue problem (10.25), deals in essence with the calculation of the transition probabilities between the unperturbed eigenstates $|\psi_n\rangle$ of the system.

10.3.1 Transition Probability

The transition probability corresponding to a transition from an initial unperturbed state $|\psi_i\rangle$ to another unperturbed state $|\psi_f\rangle$ is obtained from (10.37):

$$P_{if}(t) = \left|\langle\psi_f \mid \hat{U}_I(t,t_i) \mid \psi_i\rangle\right|^2 = \left|\langle\psi_f \mid \psi_i\rangle - \frac{i}{\hbar}\int_0^t e^{i\omega_{fi}t'}\langle\psi_f \mid \hat{V}(t') \mid \psi_i\rangle\, dt'\right.$$
$$\left. + \left(-\frac{i}{\hbar}\right)^2 \sum_n \int_0^t e^{i\omega_{fn}t_1}\langle\psi_f \mid \hat{V}(t_1) \mid \psi_n\rangle\, dt_1 \int_0^{t_1} e^{i\omega_{ni}t_2}\langle\psi_n \mid \hat{V}(t_2) \mid \psi_i\rangle\, dt_2 + \cdots\right|^2, \tag{10.38}$$

where we have used the fact that

$$\langle\psi_f \mid \hat{V}_I(t') \mid \psi_i\rangle = \langle\psi_f \mid e^{iH_0t'/\hbar}\hat{V}(t')e^{-iH_0t'/\hbar} \mid \psi_i\rangle = \langle\psi_f \mid V(t') \mid \psi_i\rangle \exp\left(i\omega_{fi}t'\right), \tag{10.39}$$

where ω_{fi} is the transition frequency between the initial and final levels i and f:

$$\omega_{fi} = \frac{E_f - E_i}{\hbar} = \frac{1}{\hbar}\left(\langle\psi_f \mid \hat{H}_0 \mid \psi_f\rangle - \langle\psi_i \mid \hat{H}_0 \mid \psi_i\rangle\right). \tag{10.40}$$

The transition probability (10.38) can be written in terms of the expansion coefficients $c_n(t)$ introduced in (10.29) as

$$P_{if}(t) = \left|c_f^{(0)} + c_f^{(1)}(t) + c_f^{(2)}(t) + \cdots\right|^2, \tag{10.41}$$

where

$$c_f^{(0)} = \langle\psi_f \mid \psi_i\rangle = \delta_{f,i}, \qquad c_f^{(1)}(t) = -\frac{i}{\hbar}\int_0^t \langle\psi_f \mid \hat{V}(t') \mid \psi_i\rangle e^{i\omega_{fi}t'}dt', \ldots. \tag{10.42}$$

The first-order transition probability for $|\psi_i\rangle \rightarrow |\psi_f\rangle$ with $i \neq f$ (and hence $\langle\psi_f \mid \psi_i\rangle = 0$) is obtained by terminating (10.38) at the first order in $\hat{V}_I(t)$:

$$\boxed{P_{if}(t) = \left|-\frac{i}{\hbar}\int_0^t \langle\psi_f \mid \hat{V}(t') \mid \psi_i\rangle e^{i\omega_{fi}t'}dt'\right|^2}. \tag{10.43}$$

In principle we can use (10.38) to calculate the transition probability to any order in $\hat{V}_I(t)$. However, terms higher than the first order become rapidly intractable. For most problems of atomic and nuclear physics, the first order (10.43) is usually sufficient. In what follows, we are going to apply (10.43) to calculate the transition probability for two cases, which will have later usefulness when we deal with the interaction of atoms with radiation: a *constant* perturbation and a *harmonic* perturbation.

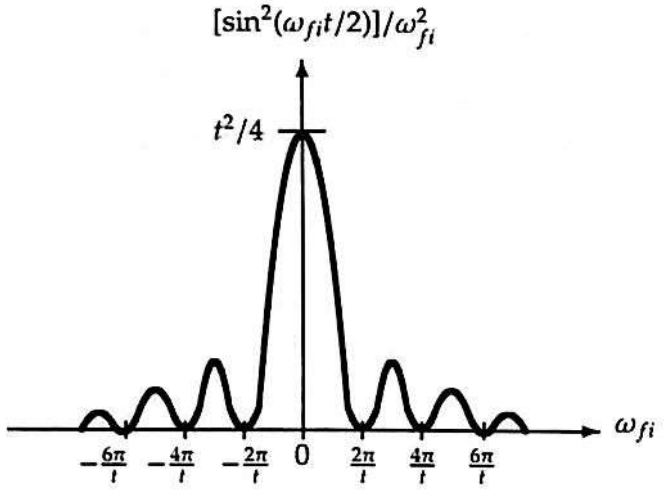


Figure 10.1: Plot of $[\sin^2(\omega_{fi}t/2)]/\omega_{fi}^2$ versus ω_{fi} for a fixed value of t; $\omega_{fi} = (E_f - E_i)/2$.

10.3.2 Transition Probability for a Constant Perturbation

In the case where $\hat{V}$ does not depend on time, (10.43) leads to

$$P_{if}(t) = \frac{1}{\hbar^2}\left|\langle\psi_f \mid \hat{V} \mid \psi_i\rangle \int_0^t e^{i\omega_{fi}t'}dt'\right|^2 = \frac{1}{\hbar^2}\left|\langle\psi_f \mid \hat{V} \mid \psi_i\rangle\right|^2 \left|\frac{e^{i\omega_{fi}t} - 1}{\omega_{fi}}\right|^2, \tag{10.44}$$

which, using $|e^{i\theta} - 1|^2 = 4\sin^2(\theta/2)$, reduces to

$$\boxed{P_{if}(t) = \frac{4\left|\langle\psi_f \mid \hat{V} \mid \psi_i\rangle\right|^2}{\hbar^2\omega_{fi}^2}\sin^2\left(\frac{\omega_{fi}t}{2}\right).} \tag{10.45}$$

As a function of time, this transition probability is an oscillating sinusoidal function with a period of $2\pi/\omega_{fi}$. As a function of ω_{fi}, however, the transition probability, as shown in Figure 10.1, has an interference pattern: it is appreciable only near $\omega_{fi} \simeq 0$ and decays rapidly as ω_{fi} moves away from zero (here, for a fixed t, we have assumed that ω_{fi} is a continuous variable; that is, we have considered a continuum of final states; we will deal with this in more detail in a moment). This means that the transition probability of finding the system in a state $|\psi_f\rangle$ of energy E_f is greatest only when $E_i \simeq E_f$ or when $\omega_{fi} \simeq 0$. The height and the width of the main peak, centered around $\omega_{fi} = 0$, are proportional to t^2 and $1/t$, respectively, so the area under the curve is proportional to t; since most of the area is under the central peak, the transition probability is proportional to t. The transition probability therefore grows linearly with time. The central peak becomes narrower and higher as time increases; this is exactly the property of a delta function. Thus, in the limit $t \to \infty$ the transition probability takes the shape of a delta function, as we are going to see.

As $t \to \infty$ we can use the asymptotic relation (Appendix A)

$$\lim_{t\to\infty} \frac{\sin^2(yt)}{\pi y^2 t} = \delta(y) \tag{10.46}$$

to write the following expression:

$$\frac{1}{(\frac{1}{2}\omega_{fi})^2}\sin^2\left(\frac{\omega_{fi}t}{2}\right) = 2\pi t\hbar\delta(\hbar\omega_{fi}), \tag{10.47}$$

because $\delta(\omega_{fi}/2) = 2\hbar\delta(\hbar\omega_{fi})$. Now since $\hbar\omega_{fi} = E_f - E_i$ and hence $\delta(\hbar\omega_{fi}) = \delta(E_f - E_i)$, we can reduce (10.45) in the limit of long times to

$$\boxed{P_{if}(t) = \frac{2\pi t}{\hbar}\left|\langle\psi_f \mid \hat{V} \mid \psi_i\rangle\right|^2 \delta(E_f - E_i).} \tag{10.48}$$

The *transition rate*, which is defined as a transition probability per unit time, is thus given by

$$\boxed{\Gamma_{if} = \frac{P_{if}(t)}{t} = \frac{2\pi}{\hbar}\left|\langle\psi_f \mid \hat{V} \mid \psi_i\rangle\right|^2 \delta(E_f - E_i).} \tag{10.49}$$

The delta term $\delta(E_f - E_i)$ guarantees the conservation of energy: in the limit $t \to \infty$, the transition rate is nonvanishing only between states of equal energy. Hence a constant (time-independent) perturbation neither removes energy from the system nor supplies energy to it. It simply causes energy-conserving transitions.

Transition into a continuum of final states

Let us now calculate the total transition rate associated with a transition from an initial state $|\psi_i\rangle$ into a continuum of final states $|\psi_f\rangle$. If $\rho(E_f)$ is the density of final states—the number of states per unit energy interval—the number of final states within the energy interval E_f and $E_f + dE_f$ is equal to $\rho(E_f)\,dE_f$. The total transition rate W_{if} can then be obtained from (10.49):

$$W_{if} = \int \frac{P_{if}(t)}{t}\rho(E_f)\,dE_f = \frac{2\pi}{\hbar}|\langle\psi_f \mid \hat{V} \mid \psi_i\rangle|^2 \int \rho(E_f)\delta(E_f - E_i)\,dE_f, \tag{10.50}$$

or

$$\boxed{W_{if} = \frac{2\pi}{\hbar}\left|\langle\psi_f \mid \hat{V} \mid \psi_i\rangle\right|^2 \rho(E_i).} \tag{10.51}$$

This relation is called the *Fermi golden rule*. It implies that, in the case of a constant perturbation, if we wait long enough, the total transition rate becomes constant (time independent).

10.3.3 Transition Probability for a Harmonic Perturbation

Consider now a perturbation which depends harmonically on time (i.e., the time between the moments of turning the perturbation on and off):

$$\hat{V}(t) = \hat{v}e^{i\omega t} + \hat{v}^\dagger e^{-i\omega t}, \tag{10.52}$$

where $\hat{v}$ is a time-independent operator. Such a perturbation is encountered, for instance, when charged particles (e.g., electrons) interact with an electromagnetic field. This perturbation provokes transitions of the system from one stationary state to another.

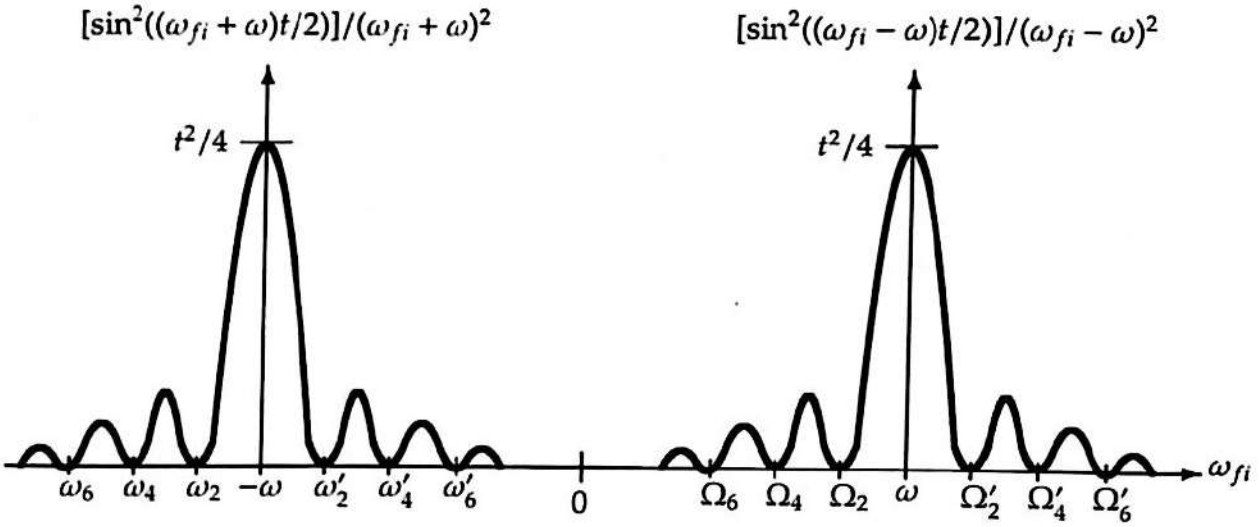


Figure 10.2: Plot of $[\sin^2((\omega_{fi} \pm \omega)t/2)]/(\omega_{fi} \pm \omega)^2$ versus ω_{fi} for a fixed value of t, where $\omega_n = -\omega - n\pi/t$, $\omega_n' = -\omega + n\pi/t$, $\Omega_n = \omega - n\pi/t$, and $\Omega_n' = \omega + n\pi/t$.

The transition probability corresponding to this perturbation can be obtained from (10.43):

$$P_{if}(t) = \frac{1}{\hbar^2}\left|\langle\psi_f \mid \hat{v} \mid \psi_i\rangle \int_0^t e^{i(\omega_{fi}+\omega)t'}dt' + \langle\psi_f \mid \hat{v}^\dagger \mid \psi_i\rangle \int_0^t e^{i(\omega_{fi}-\omega)t'}dt'\right|^2. \tag{10.53}$$

Neglecting the cross terms, for they are negligible compared with the other two (because they induce no lasting transitions), we can rewrite this expression as

$$P_{if}(t) = \frac{1}{\hbar^2}\left|\langle\psi_f \mid \hat{v} \mid \psi_i\rangle\right|^2 \left|\frac{e^{i(\omega_{fi}+\omega)t}-1}{\omega_{fi}+\omega}\right|^2 + \frac{1}{\hbar^2}\left|\langle\psi_f \mid \hat{v}^\dagger \mid \psi_i\rangle\right|^2 \left|\frac{e^{i(\omega_{fi}-\omega)t}-1}{\omega_{fi}-\omega}\right|^2, \tag{10.54}$$

which, using $|e^{i\theta}-1|^2 = 4\sin^2(\theta/2)$, reduces to

$$P_{if}(t) = \frac{4}{\hbar^2}\left[\left|\langle\psi_f \mid \hat{v} \mid \psi_i\rangle\right|^2 \frac{\sin^2((\omega_{fi}+\omega)t/2)}{(\omega_{fi}+\omega)^2} + \left|\langle\psi_f \mid \hat{v}^\dagger \mid \psi_i\rangle\right|^2 \frac{\sin^2((\omega_{fi}-\omega)t/2)}{(\omega_{fi}-\omega)^2}\right]. \tag{10.55}$$

As displayed in Figure 10.2, the transition probability peaks either at $\omega_{fi} = -\omega$, where its maximum value is $P_{if}(t) = (t^2/4\hbar^2)|\langle\psi_f \mid \hat{v} \mid \psi_i\rangle|^2$, or at $\omega_{fi} = \omega$, where its maximum value is $P_{if}(t) = (t^2/4\hbar^2)|\langle\psi_f \mid \hat{v}^\dagger \mid \psi_i\rangle|^2$. These are conditions for resonance; this means that the probability of transition is greatest only when the frequency of the perturbing field is close to $\pm\omega_{fi}$. As ω moves away from $\pm\omega_{fi}$, P_{fi} decreases rapidly.

Note that the expression (10.55) is similar to that derived for a constant perturbation, as shown in (10.45). Using (10.47) we can reduce (10.55) in the limit $t \to \infty$ to

$$\boxed{\Gamma_{if} = \frac{2\pi}{\hbar}\left|\langle\psi_f \mid \hat{v} \mid \psi_i\rangle\right|^2 \delta(E_f - E_i + \hbar\omega) + \frac{2\pi}{\hbar}\left|\langle\psi_f \mid \hat{v}^\dagger \mid \psi_i\rangle\right|^2 \delta(E_f - E_i - \hbar\omega).} \tag{10.56}$$

This transition rate is nonzero only when either of the following two conditions is satisfied:

$$E_f = E_i - \hbar\omega, \tag{10.57}$$

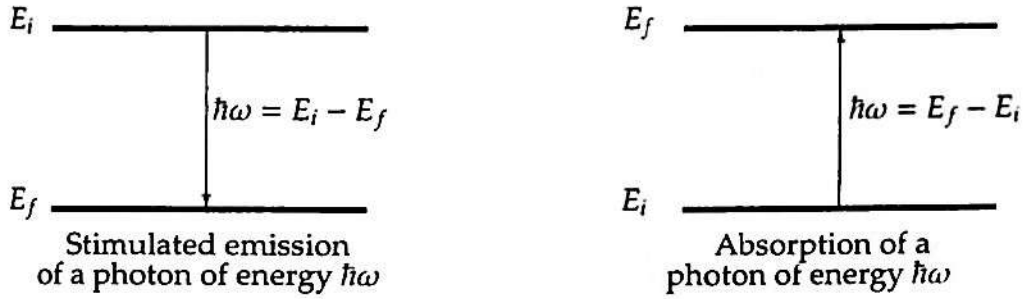


Figure 10.3: Stimulated emission and absorption of a photon of energy $\hbar\omega$.

$$E_f = E_i + \hbar\omega. \tag{10.58}$$

These two conditions cannot be satisfied simultaneously; their physical meaning can be understood as follows. The first condition $E_f = E_i - \hbar\omega$ implies that the system is initially excited, since its final energy is smaller than the initial energy; when acted upon by the perturbation, the system deexcites by giving up a photon of energy $\hbar\omega$ to the potential $\hat{V}(t)$ as shown in Figure 10.3. This process is called *stimulated emission*, since the system easily emits a photon of energy $\hbar\omega$. The second condition, $E_f = E_i + \hbar\omega$ shows that the final energy of the system is larger than its initial energy. The system then *absorbs* a photon of energy $\hbar\omega$ from $\hat{V}(t)$ and ends up in an excited state of (higher) energy E_f (Figure 10.3). We may thus view the terms $e^{i\omega t}$ and $e^{-i\omega t}$ in $\hat{V}(t)$ as responsible, respectively, for the emission and the absorption of a photon of energy $\hbar\omega$.

In conclusion, the effect of a harmonic perturbation is to transfer to the system, or to receive from it, a photon of energy $\hbar\omega$. In sharp contrast, a constant (time-independent) perturbation neither transfers energy to the system nor removes energy from it.

Remark
For transitions into a continuum of final states, we can show, by analogy with the derivation of (10.51), that (10.56) leads to the absorption and emission transition rates:

$$W_{if}^{abs} = \frac{2\pi}{\hbar} \left|\langle \psi_f \mid \hat{V}^\dagger \mid \psi_i \rangle\right|^2 \rho(E_f)\Big|_{E_f = E_i + \hbar\omega}, \tag{10.59}$$

$$W_{if}^{emi} = \frac{2\pi}{\hbar} \left|\langle \psi_f \mid \hat{V} \mid \psi_i \rangle\right|^2 \rho(E_f)\Big|_{E_f = E_i - \hbar\omega}. \tag{10.60}$$

Since the perturbation (10.52) is Hermitian, $\langle \psi_f \mid \hat{V} \mid \psi_i \rangle = \langle \psi_i | \hat{V}^\dagger \mid \psi_f \rangle^*$, we have $|\langle \psi_f \mid \hat{V} \mid \psi_i \rangle|^2 = |\langle \psi_f \mid \hat{V}^\dagger \mid \psi_i \rangle|^2$; hence

$$\frac{W_{if}^{abs}}{\rho(E_f)\big|_{E_f = E_i + \hbar\omega}} = \frac{W_{if}^{emi}}{\rho(E_f)\big|_{E_f = E_i - \hbar\omega}}. \tag{10.61}$$

This relation is known as the condition of *detailed balancing*.

Example 10.1
A particle, which is initially ($t = 0$) in the ground state of an infinite, one-dimensional potential box with walls at $x = 0$ and $x = a$, is subjected for $0 \le t \le \infty$ to a perturbation $\hat{V}(t) = \hat{x}^2 e^{-t/\tau}$. Calculate to first order the probability of finding the particle in its first excited state for $t \ge 0$.

Solution
For a particle in a box potential, with $E_n = n^2\pi^2\hbar^2/(2ma^2)$ and $\psi_n(x) = \sqrt{2/a}\sin(n\pi x/a)$, the ground state corresponds to $n = 1$ and the first excited state to $n = 2$. We can use (10.43) to obtain

$$P_{12} = \frac{1}{\hbar^2}\left| \int_0^\infty \langle \psi_2 | \hat{V}(t) | \psi_1 \rangle e^{i\omega_{21}t} dt \right|^2 = \frac{1}{\hbar^2} \left|\langle \psi_2 | \hat{x}^2 | \psi_1 \rangle\right|^2 \left| \int_0^\infty e^{-(1/\tau - i\omega_{21})t} dt \right|^2, \tag{10.62}$$

where

$$\langle \psi_2 | \hat{x}^2 | \psi_1 \rangle = \int_0^a x^2 \psi_2^*(x)\psi_1(x)\,dx = \frac{2}{a}\int_0^a x^2 \sin\left(\frac{2\pi x}{a}\right)\sin\left(\frac{\pi x}{a}\right)dx = -\frac{16a^2}{9\pi^2}, \tag{10.63}$$

$$\left| \int_0^t e^{-(1/\tau - i\omega_{21})t} dt \right|^2 = \left| \frac{e^{-(1/\tau - i\omega_{21})t} - 1}{1/\tau - i\omega_{21}} \right|^2 = \frac{1 + e^{-2t/\tau} - 2e^{-t/\tau}\cos(\omega_{21}t)}{\omega_{21}^2 + 1/\tau^2}, \tag{10.64}$$

which, in the limit $t \to \infty$, reduces to

$$\left| \int_0^\infty e^{-(1/\tau - i\omega_{21})t} dt \right|^2 = \left[\omega_{21}^2 + \frac{1}{\tau^2}\right]^{-1} = \left[\frac{9\pi^4\hbar^2}{4m^2a^4} + \frac{1}{\tau^2}\right]^{-1}, \tag{10.65}$$

since $\omega_{21} = (E_2 - E_1)/\hbar = 3\pi^2\hbar/(2ma^2)$. A substitution of (10.63) and (10.65) into (10.62) leads to

$$P_{12} = \left(\frac{16a^2}{9\pi^2\hbar}\right)^2 \left[\frac{9\pi^4\hbar^2}{4m^2a^4} + \frac{1}{\tau^2}\right]^{-1}. \tag{10.66}$$

10.4 Adiabatic and Sudden Approximations

In discussing the time-dependent perturbation theory, we have dealt with phenomena where the perturbation $\hat{V}(t)$ is small, but we have paid no attention to the rate of change of the perturbation. In this section we want to discuss approximation methods treating phenomena where $\hat{V}(t)$ is not only small but also switched on either *adiabatically* (slowly) or *suddenly* (rapidly). We assume here that $\hat{V}(t)$ is switched on at $t = 0$ and off at a later time t (the turning on and off may be smooth or abrupt).

Since $e^{i\omega_{fi}t} = (1/i\omega_{fi})\partial e^{i\omega_{fi}t}/\partial t$ an integration by parts yields

$$
\begin{aligned}
-\frac{i}{\hbar}\int_0^t \langle\psi_f \mid \hat{V}(t') \mid \psi_i\rangle e^{i\omega_{fi}t'}\,dt' &= -\frac{1}{\hbar\omega_{fi}}\int_0^t \langle\psi_f \mid \hat{V}(t') \mid \psi_i\rangle\left(\frac{\partial}{\partial t'}e^{i\omega_{fi}t'}\right)dt' \\
&= -\frac{1}{\hbar\omega_{fi}}\langle\psi_f \mid \hat{V}(t) \mid \psi_i\rangle e^{i\omega_{fi}t}\Big|_{t=0}^{t} + \frac{1}{\hbar\omega_{fi}}\int_0^t e^{i\omega_{fi}t'}\left(\frac{\partial}{\partial t'}\langle\psi_f \mid \hat{V}(t') \mid \psi_i\rangle\right)dt' \\
&= \frac{1}{\hbar\omega_{fi}}\int_0^t e^{i\omega_{fi}t'}\left(\frac{\partial}{\partial t'}\langle\psi_f \mid \hat{V}(t') \mid \psi_i\rangle\right)dt', \qquad (10.67)
\end{aligned}
$$

where we have used the fact that $\hat{V}(t)$ vanishes at the limits (when it is switched on at $t = 0$ and off at time t). The calculation of the integral depends on the rate of change of $\hat{V}(t)$. In what follows we are going to consider the cases where the interaction is switched on slowly or rapidly.

10.4.1 Adiabatic Approximation

First, let us discuss briefly the adiabatic approximation without combining it with perturbation theory. This approximation applies to phenomena whose Hamiltonians evolve *slowly* with time; we should highlight the fact that the adiabatic approximation does not require the Hamiltonian to split into an unperturbed part $\hat{H}_0$ and a weak time-dependent perturbation $\hat{V}(t)$. Essentially, it consists in approximating the solutions of the Schrödinger equation at every time by the stationary states (energy E_n and wave functions ψ_n) of the instantaneous Hamiltonian in such a way that the wave function at a given time is continuously and smoothly converted into an eigenstate of the corresponding Hamiltonian at a later time. This result is the basis of an important theorem of quantum mechanics, known as the *adiabatic theorem*, which states that: if a system is initially in the nth state and if its Hamiltonian evolves slowly with time, it will be found at a later time in the nth state of the new (instantaneous) Hamiltonian. That is, *the system will make no transitions*; it simply remains in the nth state of the new Hamiltonian.

Let us now discuss the adiabatic approximation for those cases where the Hamiltonian splits into a time-independent part $\hat{H}_0$ and a time-dependent part $\hat{V}(t)$, which is small enough so that perturbation theory applies and which is turned on and off very slowly. If $\hat{V}(t)$ is turned on at $t = 0$ and off at time t in a slow and smooth way, it will change very little in the time interval $0 \le t' \le t$. The term $\partial\langle\psi_f \mid V(t') \mid \psi_i\rangle/\partial t'$ will be almost constant, so we can take it outside the integral sign in (10.67):

$$
P_{if}(t) \simeq \frac{1}{\hbar^2\omega_{fi}^2}\left|\frac{\partial}{\partial t}\langle\psi_f \mid \hat{V}(t) \mid \psi_i\rangle\right|^2\left|\int_0^t e^{i\omega_{fi}t'}dt'\right|^2, \qquad (10.68)
$$

or

$$P_{if}(t) \simeq \frac{4}{\hbar^2 \omega_{fi}^4} \left| \frac{\partial}{\partial t} \langle \psi_f \mid \hat{V}(t) \mid \psi_i \rangle \right|^2 \sin^2\left(\frac{\omega_{fi} t}{2}\right). \tag{10.69}$$

The adiabatic approximation is valid only when the time change in the energy of the perturbation during one period of oscillation is very small compared with the energy difference $|E_f - E_i|$ between the initial and final states:

$$\left| \frac{1}{\omega_{fi}} \frac{\partial}{\partial t} \langle \psi_f \mid \hat{V}(t) \mid \psi_i \rangle \right| \ll |E_f - E_i|. \tag{10.70}$$

Since $\sin^2 \alpha \ll 1$ we see from (10.69) that, in the adiabatic approximation, the transition probability is very small, $P_{if} \ll 1$. In fact, if the rate of change of $\hat{V}(t)$, and hence of $\hat{H}(t)$, is very small, we will have $\partial \langle \psi_f \mid \hat{V}(t) \mid \psi_i \rangle / \partial t \to 0$, which in turn implies that the transition probability is practically zero: $P_{if} \to 0$. Once more, we see that *no transition occurs when the perturbation is turned on and off adiabatically.* That is, if a system is initially (at $t = 0$) in the nth state $\mid \psi_n(0) \rangle$ of $\hat{H}_0$ with energy $E_n(0)$, then at the end (at time t) of an *adiabatic perturbation* $\hat{V}(t)$, it will be found in the nth state $\mid \psi_n(t) \rangle$ of the new Hamiltonian ($\hat{H} = \hat{H}_0 + \hat{V}(t)$) with energy $E_n(t)$. As an illustrative example, consider a particle in a harmonic oscillator potential whose constant is being changed *very slowly* from k to, say, $3k$; if the particle is initially in the second excited state, it will remain in the second excited state of the new oscillator.

Note that the transition probability (10.69) was derived by making use of two approximations: the perturbation theory approximation and the adiabatic approximation. It should be stressed, however, that when the perturbation is not weak, but switched on adiabatically, we can still use the adiabatic approximation but no longer in conjunction with perturbation theory.

10.4.2 Sudden Approximation

Again, let us start with a brief discussion of the sudden approximation without invoking perturbation theory. If the Hamiltonian of a system changes abruptly (over a very short time interval) from one form to another, we would expect the wave function not to change much, yet its expansion in terms of the eigenfunctions of the initial and final Hamiltonians may be different. Consider, for instance, a system which is initially ($t < 0$) in an eigenstate $\mid \psi_n \rangle$ of the Hamiltonian $\hat{H}_0$:

$$\hat{H}_0 \mid \psi_n \rangle = E_n^{(0)} \mid \psi_n \rangle, \qquad \mid \psi_n(t) \rangle = e^{iE_n^{(0)} t/\hbar} \mid \psi_n \rangle. \tag{10.71}$$

At time $t = 0$ we assume that the Hamiltonian is suddenly changed from $\hat{H}_0$ to $\hat{H}$ and that it preserves this new form (i.e., $\hat{H}$) for $t > 0$; it should be stressed that the difference between the two Hamiltonians $\hat{H} - \hat{H}_0$ does not need to be small. Let $\mid \phi_n \rangle$ be the eigenfunctions of $\hat{H}$:

$$\hat{H} \mid \phi_n \rangle = E_n \mid \phi_n \rangle, \qquad \mid \phi_n(t) \rangle = e^{iE_n t/\hbar} \mid \phi_n \rangle. \tag{10.72}$$

The state of the system is given for $t > 0$ by

$$\mid \Phi(t) \rangle = \sum_n c_n e^{iE_n t/\hbar} \mid \phi_n \rangle. \tag{10.73}$$

If the system is initially in an eigenstate $|\psi_m\rangle$ of $\hat{H}_0$, the continuity condition at $t = 0$ dictates that the system remains in this state just after the change takes place:

$$|\Phi(0)\rangle = \sum_n c_n |\phi_n\rangle = |\psi_m\rangle \quad \Longrightarrow \quad c_n = \langle\phi_n | \psi_m\rangle. \tag{10.74}$$

The probability that a sudden change in the system's Hamiltonian from $\hat{H}_0$ to $\hat{H}$ causes a transition from the mth state of $\hat{H}_0$ to the nth state of $\hat{H}$ is

$$\boxed{P_{mn} = |\langle\phi_n | \psi_m\rangle|^2.} \tag{10.75}$$

We should note that the sudden approximation is applicable only for transitions between discrete states.

Let us now look at the sudden approximation within the context of perturbation theory. Consider a system which is subjected to a perturbation that is small and switched on suddenly. When $\hat{V}(t)$ is instantaneously turned on, the term $e^{i\omega_{fi}t'}$ in (10.67) does not change much during the switching-on time. We can therefore take $e^{i\omega_{fi}t'}$ outside the integral sign,

$$P_{if} \simeq \frac{1}{\hbar^2\omega_{fi}^2} \left| e^{i\omega_{fi}t} \right|^2 \left| \int_0^t \frac{\partial}{\partial t'} \langle\psi_f | \hat{V}(t') | \psi_i\rangle dt' \right|^2 \tag{10.76}$$

hence the transition probability is given within the sudden approximation by

$$P_{if}(t) \simeq \frac{|\langle\psi_f | \hat{V}(t) | \psi_i\rangle|^2}{\hbar^2\omega_{fi}^2}. \tag{10.77}$$

To conclude, notice that both (10.75) and (10.77) give the transition probability within the sudden approximation. Equation (10.75) represents the exact formula, where the change in the Hamiltonians, $\hat{H} - \hat{H}_0$, may be large, but equation (10.77) gives only an approximate result, for it was derived from a first-order perturbative treatment, where we assumed that the change $\hat{H} - \hat{H}_0$ is small, yet sudden.

Example 10.2

A particle is initially ($t < 0$) in the ground state of an infinite, one-dimensional potential well with walls at $x = 0$ and $x = a$.

(a) If the wall at $x = a$ is moved *slowly* to $x = 8a$, find the energy and wave function of the particle in the new well. Calculate the work done in this process.

(b) If the wall at $x = a$ is now *suddenly* moved (at $t = 0$) to $x = 8a$, calculate the probability of finding the particle in (i) the ground state, (ii) the first excited state, and (iii) the second excited state of the new potential well.

Solution

For $t < 0$ the particle was in a potential well with walls at $x = 0$ and $x = a$, and hence

$$E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}, \qquad \psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) \qquad (0 \le x \le a). \tag{10.78}$$

(a) When the wall is moved slowly, the adiabatic theorem dictates that the particle will make no transitions; it will be found at time t in the ground state of the new potential well (the well with walls at $x = 0$ and $x = 8a$). Thus, we have

$$E_1(t) = \frac{\pi^2\hbar^2}{2m(8a)^2} = \frac{\pi^2\hbar^2}{128ma^2}, \qquad \psi_1'(x) = \sqrt{\frac{2}{8a}}\sin\left(\frac{\pi x}{8a}\right) \qquad (0 \le x \le a). \tag{10.79}$$

The work needed to move the wall is

$$\Delta W = E_1(t) - E_1 = \frac{\pi^2\hbar^2}{m(8a)^2} - \frac{\pi^2\hbar^2}{2ma^2} = -\frac{63\pi^2\hbar^2}{128ma^2}. \tag{10.80}$$

(b) When the wall is moved rapidly, the particle will find itself instantly (at $t \ge 0$) in the new potential well; its energy levels and wave function are now given by

$$E_n' = \frac{n^2\pi^2\hbar^2}{2m(8a)^2} = \frac{n^2\pi^2\hbar^2}{128ma^2}, \qquad \psi_n'(x) = \sqrt{\frac{2}{8a}}\sin\left(\frac{n\pi x}{8a}\right) \qquad (0 \le x \le 8a). \tag{10.81}$$

The probability of finding the particle in the ground state of the new box potential can be obtained from (10.75): $P_{11} = |\langle\psi_1' \mid \psi_1\rangle|^2$, where

$$\langle\psi_1' \mid \psi_1\rangle = \int_0^a \psi_1'^*(x)\psi_1(x)\,dx = \frac{2}{\sqrt{8}a}\int_0^a \sin\left(\frac{\pi x}{8a}\right)\sin\left(\frac{\pi x}{a}\right)dx = \frac{16}{63\pi}\sqrt{4-2\sqrt{2}}; \tag{10.82}$$

hence

$$P_{11} = \left|\langle\psi_1' \mid \psi_1\rangle\right|^2 = \left(\frac{16}{63\pi}\right)^2(4-2\sqrt{2}) = 0.0077 \simeq 0.7\%. \tag{10.83}$$

The probability of finding the particle in the first excited state of the new box potential is given by $P_{12} = \left|\langle\psi_2' \mid \psi_1\rangle\right|^2$, where

$$\langle\psi_2' \mid \psi_1\rangle = \int_0^a \psi_2'^*(x)\psi_1(x)\,dx = \frac{2}{\sqrt{8}a}\int_0^a \sin\left(\frac{\pi x}{4a}\right)\sin\left(\frac{\pi x}{a}\right)dx = \frac{8}{15\pi}; \tag{10.84}$$

hence

$$P_{12} = \left|\langle\psi_2' \mid \psi_1\rangle\right|^2 = \left(\frac{8}{15\pi}\right)^2 = 0.1699 \simeq 17\%. \tag{10.85}$$

A similar calculation leads to

$$P_{13} = \left|\langle\psi_3' \mid \psi_1\rangle\right|^2 = \left|\frac{2}{\sqrt{8}a}\int_0^a \sin\left(\frac{3\pi x}{8a}\right)\sin\left(\frac{\pi x}{a}\right)dx\right|^2 = \left|\frac{16}{55\pi}\sqrt{4+2\sqrt{2}}\right|^2 \simeq 24.2\%. \tag{10.86}$$

These calculations show that the particle is most likely to be found in higher excited states; the probability of finding it in the ground state is very small.

10.5 Interaction of Atoms with Radiation

One of the most important applications of time-dependent perturbation theory is to study the interaction of atomic electrons with an external electromagnetic radiation. Such an application reveals a great deal about the structure of atoms. For simplicity, we assume that only one atomic electron is involved in the interaction and that the electron spin is neglected. We also assume that the nucleus is infinitely heavy.

In the absence of an external perturbation, the Hamiltonian of the atomic electron is $\hat{H}_0 = \vec{P}^2/(2m_e) + V_0(\vec{r})$, where m_e is the mass of the electron and $V_0(\vec{r})$ is the static potential due to the interaction of the electron with the other electrons and with the nucleus.

Now, if electromagnetic radiation of vector potential $\vec{A}(\vec{r},t)$ and electric potential $\phi(\vec{r},t)$ is applied on the atom, the Hamiltonian due to the interaction of the electron (of charge $-e$) with the radiation is given by

$$\begin{aligned} H &= \frac{1}{2m_e}\left(\vec{P} + \frac{e}{c}\vec{A}(\vec{r},t)\right)^2 - e\phi(\vec{r},t) + V_0(\vec{r}) \\ &= H_0 - e\phi(\vec{r},t) + \frac{e}{2m_e c}\left[2\vec{A}\cdot\vec{P} - i\hbar\vec{\nabla}\cdot\vec{A}\right] + \frac{e^2\vec{A}^2}{2m_e c^2}, \end{aligned} \tag{10.87}$$

where we have used the relation $\vec{P}\cdot\vec{A} = \vec{A}\cdot\vec{P} - i\hbar\vec{\nabla}\cdot\vec{A}$. Since $\phi(\vec{r},t) = 0$ for radiation with no electrostatic source and since $\vec{\nabla}\cdot\vec{A} = 0$ (Coulomb gauge), and neglecting the term in $\vec{A}^2$, we may write (10.87) as

$$\hat{H} = \hat{H}_0 + \frac{e}{m_e c}\vec{A}\cdot\vec{P} = \hat{H}_0 + \hat{V}(t), \tag{10.88}$$

where

$$\hat{V}(t) = \frac{e}{m_e c}\vec{A}(\vec{r},t)\cdot\vec{P}. \tag{10.89}$$

This term, which gives the interaction between the electron and the radiation, is small enough (compared to $\hat{H}_0$) to be treated by perturbation theory. We are going to use perturbation theory to study the effect of $\hat{V}(t)$ on the atom. In particular, we will focus on the transitions that are induced as a result of this perturbation.

At this level, we cannot proceed further without calculating $\vec{A}(\vec{r},t)$. In what follows, we are going to show that, using $\vec{A}(\vec{r},t)$ for an electromagnetic radiation, we obtain a $\hat{V}(t)$ which has the structure of a harmonic perturbation: $\hat{V}(t) = \hat{v}e^{-i\omega t} + \hat{v}^\dagger e^{i\omega t}$. Therefore, by analogy with a harmonic perturbation, we would expect the atom to emit or absorb photons and then undergo transitions from one state to another. For the sake of completeness, we are going to determine $\vec{A}(\vec{r},t)$ in two different ways: by treating the radiation *classically* and then *quantum mechanically*. We are going to show that, unlike a quantum treatment, a classical treatment allows only a description of stimulated emission and absorption processes, but not spontaneous emission. Spontaneous emission turns out to be a purely quantum effect.

10.5.1 Classical Treatment of the Incident Radiation

A *classical*[1] treatment of the incident radiation is valid only when *large* numbers of photons contribute to the interaction with the atom (recall that quantum mechanical effects are generally encountered only when a *finite* number of photons are involved).

From classical electrodynamics, if we consider the incident radiation to be a plane wave of polarization $\vec{\varepsilon}$ that is propagating along the direction $\vec{n}$, the vector potential $\vec{A}(\vec{r},t)$ is given by

$$\vec{A}(\vec{r},t) = \vec{A}_0(\vec{r})e^{-i\omega t} + \vec{A}_0^*(\vec{r})e^{i\omega t} = A_0\vec{\varepsilon}\left[e^{i(\vec{k}\cdot\vec{r}-\omega t)} + e^{-i(\vec{k}\cdot\vec{r}-\omega t)}\right], \tag{10.90}$$

with $\vec{k} = k\vec{n}$. Since $\vec{A}(\vec{r},t)$ satisfies the wave equation $\vec{\nabla}^2\vec{A} - (1/c^2)\partial^2\vec{A}/\partial t^2 = 0$, we have $k = \omega/c$. The Coulomb gauge condition $\vec{\nabla}\cdot\vec{A} = 0$ yields $\vec{k}\cdot\vec{A}_0 = 0$; that is, $\vec{A}(\vec{r},t)$ lies in a plane perpendicular to the wave's direction of propagation, $\vec{n}$. The electric and magnetic fields associated with the vector potential (10.90) can be obtained at once:

$$\vec{E}(\vec{r},t) = -\frac{1}{c}\frac{\partial\vec{A}}{\partial t} = \frac{i\omega}{c}A_0\vec{\varepsilon}\left[-e^{i(\vec{k}\cdot\vec{r}-\omega t)} + e^{-i(\vec{k}\cdot\vec{r}-\omega t)}\right], \tag{10.91}$$

$$\vec{B}(\vec{r},t) = \vec{\nabla}\times\vec{A} = i(\vec{k}\times\vec{\varepsilon})A_0\left[-e^{i(\vec{k}\cdot\vec{r}-\omega t)} + e^{-i(\vec{k}\cdot\vec{r}-\omega t)}\right] = \vec{n}\times\vec{E}. \tag{10.92}$$

These two relations show that $\vec{E}$ and $\vec{B}$ have the same magnitude, $|\vec{E}| = |\vec{B}|$.

The energy density (or energy per unit volume) for a single photon of the incident radiation can be obtained from (10.91) and (10.92):

$$u = \frac{1}{8\pi}(|\vec{E}|^2 + |\vec{B}|^2) = \frac{1}{4\pi}|\vec{E}|^2 = \frac{\omega^2}{\pi c^2}|A_0|^2\sin^2(\vec{k}\cdot\vec{r} - \omega t). \tag{10.93}$$

Averaging this expression over time, we see that the energy of a single photon per unit volume, $\hbar\omega/V$, is given by $(\omega^2/2\pi c^2)|A_0|^2 = \hbar\omega/V$ and hence $|A_0|^2 = 2\pi\hbar c^2/(\omega V)$, which, when inserted into (10.90), leads to

$$\vec{A}(\vec{r},t) = \sqrt{\frac{2\pi\hbar c^2}{\omega V}}\left[e^{i(\vec{k}\cdot\vec{r}-\omega t)} + e^{-i(\vec{k}\cdot\vec{r}-\omega t)}\right]\vec{\varepsilon}. \tag{10.94}$$

Having specified $\vec{A}(\vec{r},t)$ by means of a classical treatment, we can now rewrite the potential (10.89) as

$$\hat{V}(t) = \frac{e}{m_e c}\left(\frac{2\pi\hbar c^2}{\omega V}\right)^{1/2}\vec{\varepsilon}\cdot\vec{p}\left[e^{i(\vec{k}\cdot\vec{r}-\omega t)} + e^{-i(\vec{k}\cdot\vec{r}-\omega t)}\right] = \hat{v}e^{-i\omega t} + \hat{v}^\dagger e^{i\omega t}, \tag{10.95}$$

where

$$\hat{v} = \frac{e}{m_e}\left(\frac{2\pi\hbar}{\omega V}\right)^{1/2}\vec{\varepsilon}\cdot\vec{p}e^{i\vec{k}\cdot\vec{r}}, \qquad \hat{v}^\dagger = \frac{e}{m_e}\left(\frac{2\pi\hbar}{\omega V}\right)^{1/2}\vec{\varepsilon}\cdot\vec{p}e^{-i\vec{k}\cdot\vec{r}}. \tag{10.96}$$

The structure of (10.95) is identical with (10.52); that is, *the interaction of an atomic electron with radiation has the structure of a harmonic perturbation*. By analogy with (10.52) we can state that the term $e^{-i\omega t}$ in (10.95) gives rise to the *absorption* of the incident photon of energy $\hbar\omega$ by

[1] A *classical* treatment of the electric and magnetic fields, $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$, and their corresponding electric and vector potentials, $\phi(\vec{r},t)$ and $\vec{A}(\vec{r},t)$, means that they are described by *continuous* fields.

the atom, and $e^{i\omega t}$ to the *stimulated emission* of a photon of energy $\hbar\omega$ by the atom. That is, the absorption process occurs when the atom receives a photon from the radiation, and the stimulated emission when the radiation receives or gains a photon from the decaying atom. At this level, we cannot afford not to mention an important application of stimulated emission. In this process we start with one (incident) photon and end up with two: the incident photon plus the photon given by the atom resulting from its transition to a lower energy level. What would happen if we had a large number of atoms in the same excited state? A single external photon would trigger an avalanche, or chain reaction, of photons released by these atoms in a very short time and all having the same frequency. This would lead to an *amplification* of the electromagnetic field. How does this take place? When the incident photon interacts with the first atom, it will produce two photons, which in turn produce four photons; these four photons then produce eight photons (after they interact with four different atoms), and so on. This process is known as the *amplification by stimulated emission of the* (incident) *radiation*. Two such radiation amplifications have been achieved experimentally and have led to enormous applications: one in the microwave domain, known as *maser* (microwave amplification by stimulated emission of radiation); the other in the domain of light waves, called *laser* (light amplification by stimulated emission of radiation).

Following the approach that led to the transition rates (10.56) from (10.52), we can easily show that the transition rates for the stimulated emission and absorption corresponding to (10.95) are given by

$$\boxed{\Gamma_{i\to f}^{emi} = \frac{4\pi^2 e^2}{m_e^2 \omega V} \left| \langle \psi_f \mid e^{-i\vec{k}\cdot\vec{r}}\, \vec{\varepsilon}\cdot\vec{P} \mid \psi_i \rangle \right|^2 \delta(E_f - E_i + \hbar\omega),} \tag{10.97}$$

$$\boxed{\Gamma_{i\to f}^{abs} = \frac{4\pi^2 e^2}{m_e^2 \omega V} \left| \langle \psi_f \mid e^{i\vec{k}\cdot\vec{r}}\, \vec{\varepsilon}\cdot\vec{P} \mid \psi_i \rangle \right|^2 \delta(E_f - E_i - \hbar\omega).} \tag{10.98}$$

These relations represent the expressions for the transition rates when the radiation is treated classically.

What would happen when there is no radiation? If $\vec{A} = 0$ (i.e., the atom is placed in a vacuum), equations (10.97) and (10.98) imply that no transition will occur since, as equation (10.89) shows, if $\vec{A} = 0$ the perturbation will be zero; hence $\Gamma_{i\to f}^{emi} = 0$ and $\Gamma_{i\to f}^{abs} = 0$. As a result, the classical treatment cannot account for *spontaneous* emission which occurs even in the absence of an external perturbing field. This implies, for instance, that a hydrogen atom in an $n \geq 2$ energy eigenstate remains in this eigenstate unless it is perturbed by an external field. This is in complete disagreement with experimental observations, which show that atoms in the $n \geq 2$ states undergo *spontaneous emissions*; they emit electromagnetic radiation *even when no external perturbation is present*. The spontaneous emission is *a purely quantum effect*.

10.5.2 Quantization of the Electromagnetic Field

We have seen that a classical treatment of radiation leads to transition rates that account only for the processes of absorption and stimulated emission; spontaneous emission of photons by atoms is a typical phenomenon that a classical treatment fails to explain, let alone predict. The classical treatment is valid only when very large numbers of photons contribute to the radiation; that is, when the intensity of the radiation is so high that only its wave aspect is

important. At very low intensities, however, the particle nature of the radiation becomes nonnegligible. In this case we have to consider a quantum mechanical treatment of the electromagnetic radiation. To obtain a quantum description of the radiation, we would necessarily need to *replace the various fields* (such as $\vec{E}(\vec{r},t)$, $\vec{B}(\vec{r},t)$ *and the potential vector* $\vec{A}(\vec{r},t)$) *with operators.*

In the absence of charges and currents, the electric and magnetic fields are fully specified by the vector potential $\vec{A}(\vec{r},t)$. Since $\vec{A}(\vec{r},t)$ is transverse (perpendicular to the wave vector $\vec{k}$), it has only two nonzero components along the directions of two polarization (unit) vectors, $\vec{\varepsilon}_1$ and $\vec{\varepsilon}_2$, which lie in a plane perpendicular to $\vec{k}$. We can thus expand $\vec{A}(\vec{r},t)$ in a Fourier series as follows:

$$\vec{A}(\vec{r},t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} \sum_{\lambda=1}^{2} \left[A_{\lambda,\vec{k}} \vec{\varepsilon}_\lambda e^{i(\vec{k}\cdot\vec{r}-\omega_k t)} + A^*_{\lambda,\vec{k}} \vec{\varepsilon}^*_\lambda e^{-i(\vec{k}\cdot\vec{r}-\omega_k t)} \right], \tag{10.99}$$

where we have assumed that the electromagnetic field is confined to a large volume V with periodic boundary conditions. We are going to see that, by analogy with the quantization of a classical harmonic oscillator, *the quantization of radiation can be achieved by writing the electromagnetic field in terms of creation and annihilation operators.*

The Hamiltonian of the complete system (atom and the external radiation) is

$$\hat{H} = \hat{H}_0 + \hat{H}_r + \hat{V}(t), \tag{10.100}$$

where $\hat{H}_0$ is the Hamiltonian of the unperturbed atom, $\hat{H}_r$ is the Hamiltonian of the electromagnetic field, and $\hat{V}(t)$ is the interaction of the atom with the radiation. To find $\hat{H}_r$ we need to quantize the energy of the electromagnetic field which can be obtained from (10.99):

$$\hat{H}_r = \frac{1}{8\pi} \int d^3r \left[\vec{E}^2(\vec{r},t) + \vec{B}^2(\vec{r},t) \right] = \frac{V}{8\pi c^2} \sum_{\vec{k}} \sum_{\lambda=1}^{2} (\hbar\vec{k})^2 A^*_{\lambda,\vec{k}} A_{\lambda,\vec{k}}, \tag{10.101}$$

with $|\vec{\varepsilon}_\lambda|^2 = 1$, where we have used $\omega_k = ck$, $\vec{E}(\vec{r},t) = -(1/c)\partial\vec{A}/\partial t$, and $\vec{B}(\vec{r},t) = \vec{\nabla} \times \vec{A}$. Instead of the two variables $A_{\lambda,\vec{k}}$ and $A^*_{\lambda,\vec{k}}$, we can introduce a new set of two canonically conjugate variables:

$$Q_{\lambda,\vec{k}} = \frac{1}{\sqrt{4\pi c^2}} \left(A^*_{\lambda,\vec{k}} + A_{\lambda,\vec{k}} \right), \qquad P_{\lambda,\vec{k}} = \frac{i\omega_k}{\sqrt{4\pi c^2}} \left(A^*_{\lambda,\vec{k}} - A_{\lambda,\vec{k}} \right). \tag{10.102}$$

Combining (10.101) and (10.102), we can write

$$H_r = \sum_{\vec{k}} \sum_{\lambda=1}^{2} \left(\frac{1}{2} P^2_{\lambda,\vec{k}} + \frac{\omega_k^2}{2} Q^2_{\lambda,\vec{k}} \right). \tag{10.103}$$

This expression has the structure of a Hamiltonian of a collection of independent harmonic oscillators. This is compatible with the fact that electromagnetic waves in a vacuum result from the (harmonic) oscillations of the electromagnetic field; hence they can be described by means of a linear superposition of independent vibrational modes. To quantize (10.103) we simply need to find the operators $\hat{Q}_{\lambda,\vec{k}}$ and $\hat{P}_{\lambda,\vec{k}}$ that correspond to the variables $Q_{\lambda,\vec{k}}$ and $P_{\lambda,\vec{k}}$, respectively, such that they obey the canonical commutation relations:

$$\left[\hat{Q}_{\lambda_1,\vec{k}_1}, \hat{P}_{\lambda_2,\vec{k}_2} \right] = i\hbar\delta_{\lambda_1,\lambda_2}\delta_{\vec{k}_1,\vec{k}_2}. \tag{10.104}$$

Following the same quantization procedure of a classical harmonic oscillator, and introducing the lowering and raising operators

$$\hat{a}_{\lambda,\vec{k}} = \sqrt{\frac{\omega_k}{2\hbar}}\hat{Q}_{\lambda,\vec{k}} + \frac{i}{\sqrt{2\hbar\omega_k}}\hat{P}_{\lambda,\vec{k}}, \qquad \hat{a}^{\dagger}_{\lambda,\vec{k}} = \sqrt{\frac{\omega_k}{2\hbar}}\hat{Q}_{\lambda,\vec{k}} - \frac{i}{\sqrt{2\hbar\omega_k}}\hat{P}_{\lambda,\vec{k}}, \tag{10.105}$$

which lead to $\hat{Q}_{\lambda,\vec{k}} = \sqrt{\hbar/2\omega_k}(\hat{a}^{\dagger}_{\lambda,\vec{k}} + \hat{a}_{\lambda,\vec{k}})$ and $\hat{P}_{\lambda,\vec{k}} = i\sqrt{\hbar\omega_k/2}(\hat{a}^{\dagger}_{\lambda,\vec{k}} - \hat{a}_{\lambda,\vec{k}})$, we can show that the Hamiltonian operator corresponding to (10.103) is given by

$$\hat{H}_r = \sum_{\vec{k}}\sum_{\lambda=1}^{2}\hbar\omega_k\left(\hat{N}_{\lambda,\vec{k}} + \frac{1}{2}\right), \tag{10.106}$$

with $\hat{N}_{\lambda,\vec{k}} = \hat{a}^{\dagger}_{\lambda,\vec{k}}\hat{a}_{\lambda,\vec{k}}$.

By analogy to the harmonic oscillator, the operators $\hat{a}_{\lambda,\vec{k}}$ and $\hat{a}^{\dagger}_{\lambda,\vec{k}}$ obey the following commutation relations:

$$\left[\hat{a}_{\lambda_1,\vec{k}_1}, \hat{a}^{\dagger}_{\lambda_2,\vec{k}_2}\right] = \delta_{\lambda_1,\lambda_2}\delta_{\vec{k}_1,\vec{k}_2}, \qquad \left[\hat{a}_{\lambda_1,\vec{k}_1}, \hat{a}_{\lambda_2,\vec{k}_2}\right] = \left[\hat{a}^{\dagger}_{\lambda_1,\vec{k}_1}, \hat{a}^{\dagger}_{\lambda_2,\vec{k}_2}\right] = 0, \tag{10.107}$$

and serve respectively to *annihilate* and *create* a photon of wave number $\vec{k}$ and polarization λ. The eigenvalues of $\hat{N}_{\lambda,\vec{k}}$ are $n_{\lambda,\vec{k}} = 0,1,2,\ldots$; by analogy with the harmonic oscillator, its eigenvectors are

$$|\, n_{\lambda,\vec{k}}\rangle = \frac{1}{\sqrt{n_{\lambda,\vec{k}}!}}\left(\hat{a}^{\dagger}_{\lambda\vec{k}}\right)^{n_{\lambda,\vec{k}}}\,|\,0\rangle, \tag{10.108}$$

where $|\,0\rangle$ is the state with no photons, the vacuum state, and $|\,n_{\lambda,\vec{k}}\rangle$ is a state of the electromagnetic field with $n_{\lambda,\vec{k}}$ photons with wave vector $\vec{k}$ and polarization λ. The number $n_{\lambda,\vec{k}}$ therefore represents the *occupation number* mode. The actions of $\hat{a}_{\lambda,\vec{k}}$ and $\hat{a}^{\dagger}_{\lambda,\vec{k}}$ on $|\, n_{\lambda,\vec{k}}\rangle$ are given by

$$\hat{a}_{\lambda,\vec{k}}\,|\, n_{\lambda,\vec{k}}\rangle = \sqrt{n_{\lambda,\vec{k}}}\,|\, n_{\lambda,\vec{k}} - 1\rangle, \qquad \hat{a}^{\dagger}_{\lambda,\vec{k}}\,|\, n_{\lambda,\vec{k}}\rangle = \sqrt{n_{\lambda,\vec{k}} + 1}\,|\, n_{\lambda,\vec{k}} + 1\rangle. \tag{10.109}$$

The eigenstates of the Hamiltonian (10.106) can be inferred from (10.108):

$$|n_{\lambda_1\vec{k}_1}, n_{\lambda_2\vec{k}_2}, n_{\lambda_3\vec{k}_3},\ldots\rangle = \prod_j |n_{\lambda_j\vec{k}_j}\rangle, \tag{10.110}$$

with the energy eigenvalues (of the radiation)

$$E_r = \sum_{\vec{k}}\sum_{\lambda}\hbar\omega_k\left(n_{\lambda\vec{k}} + \frac{1}{2}\right). \tag{10.111}$$

The state $|n_{\lambda_1\vec{k}_1}, n_{\lambda_2\vec{k}_2}, n_{\lambda_3\vec{k}_3},\ldots\rangle$ describes an electromagnetic field with $n_{\lambda_1\vec{k}_1}$ photons in the mode $(\lambda_1,\vec{k}_1)$ (i.e., $n_{\lambda_1\vec{k}_1}$ photons with wave vector $\vec{k}_1$ and polarization λ_1), $n_{\lambda_2\vec{k}_2}$ photons in the

mode $(\lambda_2, \vec{k}_2)$, and so on. Substituting (10.102) into (10.105), we get $\hat{a}_{\lambda,\vec{k}} = \sqrt{\omega_k/(2\pi\hbar c^2)}\hat{A}_{\lambda,\vec{k}}$ and $\hat{a}^\dagger_{\lambda,\vec{k}} = \sqrt{\omega_k/(2\pi\hbar c^2)}\hat{A}^\dagger_{\lambda,\vec{k}}$; hence

$$\hat{A}_{\lambda,\vec{k}} = \sqrt{\frac{2\pi\hbar c^2}{\omega_k}}\hat{a}_{\lambda,\vec{k}}, \qquad \hat{A}^\dagger_{\lambda,\vec{k}} = \sqrt{\frac{2\pi\hbar c^2}{\omega_k}}\hat{a}^\dagger_{\lambda,\vec{k}}. \tag{10.112}$$

An insertion of these two relations into (10.99) gives the vector potential operator:

$$\hat{\vec{A}}(\vec{r},t) = \sum_{\vec{k}}\sum_{\lambda=1}^{2}\sqrt{\frac{2\pi\hbar c^2}{\omega_k V}}\left[\hat{a}_{\lambda,\vec{k}}e^{i(\vec{k}\cdot\vec{r}-\omega_k t)}\,\vec{\varepsilon}_\lambda + \hat{a}^\dagger_{\lambda,\vec{k}}e^{-i(\vec{k}\cdot\vec{r}-\omega_k t)}\,\vec{\varepsilon}^{\,*}_\lambda\right]. \tag{10.113}$$

The interaction $\hat{V}(t)$ as given by (10.89) reduces to $\hat{V}(t) = (e/m_e c)\hat{\vec{A}}(\vec{r},t)\cdot\vec{P}$ or

$$\hat{V}(t) = \frac{e}{m_e}\sum_{\vec{k}}\sum_{\lambda}\sqrt{\frac{2\pi\hbar}{\omega_k V}}\left[\hat{a}_{\lambda,\vec{k}}e^{i\vec{k}\cdot\vec{r}}\vec{\varepsilon}_\lambda\cdot\vec{P}e^{i\omega_k t} + \hat{a}^\dagger_{\lambda,\vec{k}}e^{i\vec{k}\cdot\vec{r}}\vec{\varepsilon}^{\,*}_\lambda\cdot\vec{P}e^{-i\omega_k t}\right], \tag{10.114}$$

or

$$\hat{V}(t) = \sum_{\vec{k}}\sum_{\lambda=1}^{2}\left(\hat{v}_{\lambda,\vec{k}}e^{i\omega_k t} + \hat{v}^\dagger_{\lambda,\vec{k}}e^{-i\omega_k t}\right), \tag{10.115}$$

where

$$\hat{v}_{\lambda,\vec{k}} = \frac{e}{m_e}\sqrt{\frac{2\pi\hbar}{\omega_k V}}\hat{a}_{\lambda,\vec{k}}e^{i\vec{k}\cdot\vec{r}}\vec{\varepsilon}_\lambda\cdot\vec{P}, \qquad \hat{v}^\dagger_{\lambda,\vec{k}} = \frac{e}{m_e}\sqrt{\frac{2\pi\hbar}{\omega_k V}}\hat{a}^\dagger_{\lambda,\vec{k}}e^{-i\vec{k}\cdot\vec{r}}\vec{\varepsilon}^{\,*}_\lambda\cdot\vec{P}. \tag{10.116}$$

The terms $\hat{v}_{\lambda,\vec{k}}$ and $\hat{v}^\dagger_{\lambda,\vec{k}}$ correspond to the absorption (annihilation) and emission (creation) of a photon by the atom, respectively. As in the classical case, the interaction (10.115) has the structure of a *harmonic* perturbation.

Remark

The quantization of the radiation is achieved by writing the electromagnetic field in terms of creation and annihilation operators, by analogy with the harmonic oscillator. This process, which is called *second quantization*, leads to the replacement of the various *fields* (such as the vector potential $\vec{A}(\vec{r},t)$, the electric field $\vec{E}(\vec{r},t)$, and the magnetic field $\vec{B}(\vec{r},t)$) by operator quantities, which in turn are expressed in terms of creation and annihilation operators. For instance, the Hamiltonian and the vector potential of the radiation are given in the second quantization representation by equations (10.106) and (10.113), respectively.

10.5.3 Transition Rates for Absorption and Emission of Radiation

Before the atom and the radiation interact, their initial state is given by $|\,\Phi_i\rangle = |\,\psi_i\rangle\,|\,n_{\lambda,\vec{k}}\rangle$, where $|\,\psi_i\rangle$ is the state of the unperturbed atom and $|\,n_{\lambda,\vec{k}}\rangle$ is the state vector of the radiation. After the interaction takes place, the state of the system is given by $|\,\Phi_f\rangle = |\,\psi_f\rangle\,|\,n_{\lambda,\vec{k}}\rangle_f$.

Let us look first at the case of emission of a photon. If after interaction the atom emits a photon, the final state of the system will be given by $|\,\Phi_f\rangle = |\,\psi_f\rangle\,|\,n_{\lambda,\vec{k}}+1\rangle$, since the

electromagnetic field gains a photon; hence its state changes from $|\, n_{\lambda\vec{k}}\rangle \rightarrow |\, n_{\lambda\vec{k}}+1\rangle$. Formally, this process can be achieved by creating a photon, that is, by applying $\hat{v}^{\dagger}_{\lambda\vec{k}}$ or $\hat{a}^{\dagger}_{\lambda\vec{k}}$ on the photonic state $|\, n_{\lambda\vec{k}}\rangle$:

$$\begin{aligned}\langle \Phi_f \,|\, \hat{v}^{\dagger}_{\lambda\vec{k}} \,|\, \Phi_i\rangle &= \frac{e}{m_e}\sqrt{\frac{2\pi\hbar}{\omega_k V}}\langle \psi_f \,|\, e^{-i\vec{k}\cdot\vec{r}}\vec{\varepsilon}^{\,*}_{\lambda}\cdot\hat{\vec{P}} \,|\, \psi_i\rangle\langle n_{\lambda\vec{k}}+1|\hat{a}^{\dagger}_{\lambda\vec{k}}|n_{\lambda\vec{k}}\rangle \\ &= \frac{e}{m_e}\sqrt{\frac{2\pi\hbar}{\omega_k V}}\sqrt{n_{\lambda\vec{k}}+1}\langle \psi_f \,|\, e^{-i\vec{k}\cdot\vec{r}}\vec{\varepsilon}^{\,*}_{\lambda}\cdot\hat{\vec{P}} \,|\, \psi_i\rangle. \end{aligned} \tag{10.117}$$

When $n_{\lambda\vec{k}} = 0$ (i.e., no radiation), equation (10.117) shows that *even in the absence of an external radiation*, the theory can describe events where there is emission of a photon. This is called *spontaneous emission*. This phenomenon cannot be described by means of a classical treatment of radiation. But if $n_{\lambda\vec{k}} \neq 0$, then $n_{\lambda\vec{k}}$ is responsible for *induced* or *stimulated* emissions; the bigger $n_{\lambda\vec{k}}$, the bigger the emission probability.

In the case of a photon absorption, the system undergoes a transition from an initial state $|\, \Phi_i\rangle = |\, \psi_i\rangle \,|\, n_{\lambda\vec{k}}\rangle$ to the final state $|\, \Phi_f\rangle = |\, \psi_f\rangle \,|\, n_{\lambda\vec{k}} - 1\rangle$. This can be achieved formally by applying the annihilation operator $\hat{a}_{\lambda\vec{k}}$ on $|\, n_{\lambda\vec{k}}\rangle$:

$$\begin{aligned}\langle \Phi_f \,|\, \hat{v}_{\lambda\vec{k}} \,|\, \Phi_i\rangle &= \frac{e}{m_e}\sqrt{\frac{2\pi\hbar}{\omega_k V}}\langle \psi_f \,|\, e^{i\vec{k}\cdot\vec{r}}\vec{\varepsilon}_{\lambda}\cdot\hat{\vec{P}} \,|\, \psi_i\rangle\langle n_{\lambda\vec{k}}-1|\hat{a}_{\lambda\vec{k}} \,|\, n_{\lambda\vec{k}}\rangle \\ &= \frac{e}{m_e}\sqrt{\frac{2\pi\hbar}{\omega_k V}}\sqrt{n_{\lambda\vec{k}}}\langle \psi_f \,|\, e^{i\vec{k}\cdot\vec{r}}\vec{\varepsilon}_{\lambda}\cdot\hat{\vec{P}} \,|\, \psi_i\rangle. \end{aligned} \tag{10.118}$$

The transition rates corresponding to the emission or absorption of a photon of energy $\hbar\omega_k = \hbar ck$, wave number $\vec{k}$, and polarization λ can be obtained, by analogy with (10.97) and (10.98), from (10.117) and (10.118):

$$\boxed{\Gamma^{emi}_{i\to f} = \frac{4\pi^2 e^2}{m_e^2\omega_k V}\left(n_{\lambda\vec{k}}+1\right)\left|\langle \psi_f \,|\, e^{-i\vec{k}\cdot\vec{r}}\vec{\varepsilon}^{\,*}_{\lambda}\cdot\vec{P} \,|\, \psi_i\rangle\right|^2 \delta(E_f - E_i + \hbar\omega_k),} \tag{10.119}$$

$$\boxed{\Gamma^{abs}_{i\to f} = \frac{4\pi^2 e^2}{m_e^2\omega_k V}n_{\lambda\vec{k}}\left|\langle \psi_f \,|\, e^{i\vec{k}\cdot\vec{r}}\vec{\varepsilon}_{\lambda}\cdot\vec{P} \,|\, \psi_i\rangle\right|^2 \delta(E_f - E_i - \hbar\omega_k).} \tag{10.120}$$

10.5.4 Transition Rates within the Dipole Approximation

Approximate expressions of the transition rates (10.119) and (10.120) can be obtained by expanding $e^{\pm i\vec{k}\cdot\vec{r}}$:

$$e^{\pm i\vec{k}\cdot\vec{r}} = 1 \pm i\vec{k}\cdot\vec{r} - \frac{1}{2}(\vec{k}\cdot\vec{r}) \mp \cdots = 1 \pm i\frac{\omega}{c}\vec{n}\cdot\vec{r} - \frac{1}{2}\frac{\omega^2}{c^2}(\vec{n}\cdot\vec{r})^2 \mp \cdots. \tag{10.121}$$

This expansion finds its justification in the fact that $(\vec{k}\cdot\vec{r})$ is a small quantity, since the wavelength of the radiation (visible or ultraviolet) is very large compared to the atomic size: $kr = 2\pi a_0/\lambda \sim 2\pi \times 10^{-10}\,\text{m}/10^{-6}\,\text{m} \sim 10^{-3}$. In the case of nuclear radiation (such as γ radiation), kr is also in the range of 10^{-3}, with $r_{nucleus} \sim 10^{-15}\,\text{m}$.

The *electric dipole approximation* corresponds to keeping only the leading term in the expansion (10.121): $e^{\pm i\vec{k}\cdot\vec{r}} \simeq 1$; hence

$$\langle \psi_f \mid e^{\pm i\vec{k}\cdot\vec{r}}\vec{\varepsilon}_\lambda \cdot \vec{P} \mid \psi_i\rangle \simeq \vec{\varepsilon}_\lambda \cdot \langle \psi_f \mid \vec{P} \mid \psi_i\rangle. \tag{10.122}$$

This term gives rise to *electric dipole* or E1 transitions. To calculate this term, we need to use the relation

$$\left[\hat{X},\, \hat{H}_0\right] = \left[\hat{X},\, \frac{\hat{P}^2}{2m_e} + \hat{V}(\vec{r})\right] = \left[\hat{X},\, \frac{\hat{P}_x^2}{2m_e}\right] = \frac{i\hbar}{m_e}\hat{P}_x, \tag{10.123}$$

which can be generalized to $[\vec{r},\, \hat{H}_0] = i\hbar\hat{\vec{P}}/m_e$. Hence, inserting $\hat{\vec{P}} = (m_e/i\hbar)[\vec{r},\, \hat{H}_0]$ into (10.122) and using $\hat{H}_0 \mid \psi_i\rangle = E_i \mid \psi_i\rangle$ and $\hat{H}_0 \mid \psi_f\rangle = E_f \mid \psi_f\rangle$, we have

$$\begin{aligned}\vec{\varepsilon}_\lambda \cdot \langle \psi_f \mid \vec{P} \mid \psi_i\rangle &= \frac{m_e}{i\hbar}\vec{\varepsilon}_\lambda \cdot \langle \psi_f \mid [\vec{r}, \hat{H}_0] \mid \psi_i\rangle = \frac{m}{i\hbar}(E_i - E_f)\vec{\varepsilon}_\lambda \cdot \langle \psi_f \mid \vec{r} \mid \psi_i\rangle \\ &= im_e\omega_{fi}\vec{\varepsilon}_\lambda \cdot \langle \psi_f \mid \vec{r} \mid \psi_i\rangle.\end{aligned} \tag{10.124}$$

The substitution of this term into (10.122) leads to

$$\boxed{\langle \psi_f \mid e^{i\vec{k}\cdot\vec{r}}\vec{\varepsilon}_\lambda \cdot \vec{P} \mid \psi_i\rangle = im_e\omega_{fi}\vec{\varepsilon}_\lambda \cdot \langle \psi_f \mid \vec{r} \mid \psi_i\rangle.} \tag{10.125}$$

Inserting (10.125) into (10.119) and (10.120), we obtain the transition rates, within the *dipole approximation*, for the emission and absorption of a photon of energy $\hbar\omega_k$ by the atom:

$$\boxed{\Gamma_{i\to f}^{emi} = \frac{4\pi^2 e^2 \omega_{fi}^2}{\omega_k V}(n_{\lambda,k} + 1)\left|\vec{\varepsilon}_\lambda^{\,*} \cdot \langle \psi_f \mid \vec{r} \mid \psi_i\rangle\right|^2 \delta(E_f - E_i + \hbar\omega_k),} \tag{10.126}$$

$$\boxed{\Gamma_{i\to f}^{abs} = \frac{4\pi^2 e^2 \omega_{fi}^2}{\omega_k V} n_{\lambda,k}\left|\vec{\varepsilon}_\lambda \cdot \langle \psi_f \mid \vec{r} \mid \psi_i\rangle\right|^2 \delta(E_f - E_i - \hbar\omega_k).} \tag{10.127}$$

10.5.5 The Electric Dipole Selection Rules

Since $\vec{r}$ is given in spherical coordinates by $\vec{r} = (r\sin\theta\cos\phi)\hat{\imath} + (r\sin\theta\sin\phi)\hat{\jmath} + (r\cos\theta)\vec{k}$, we can write

$$\vec{\varepsilon}_\lambda \cdot \vec{r} = r(\varepsilon_x \sin\theta\cos\phi + \varepsilon_y \sin\theta\sin\phi + \varepsilon_z\cos\theta). \tag{10.128}$$

Using the relations $\sin\theta\cos\phi = -\sqrt{2\pi/3}(Y_{11} - Y_{1-1})$, $\sin\theta\sin\phi = i\sqrt{2\pi/3}(Y_{11} + Y_{1-1})$, and $\cos\theta = \sqrt{4\pi/3}\,Y_{10}$, we may rewrite (10.128) as

$$\vec{\varepsilon}_\lambda \cdot \vec{r} = \sqrt{\frac{4\pi}{3}}r\left(\frac{-\varepsilon_x + i\varepsilon_y}{\sqrt{2}}Y_{11} + \frac{\varepsilon_x + i\varepsilon_y}{\sqrt{2}}Y_{1-1} + \varepsilon_z Y_{10}\right), \tag{10.129}$$

which in turn leads to

$$\begin{aligned}\langle \psi_f \mid \vec{\varepsilon}_\lambda \cdot \vec{r} \mid \psi_i\rangle &= \sqrt{\frac{4\pi}{3}}\int_0^\infty r^3 R^*_{n_f l_f}(r)R_{n_i l_i}(r)\,dr \\ &\times \int Y^*_{l_f m_f}(\theta,\varphi)\left(\frac{-\varepsilon_x + i\varepsilon_y}{\sqrt{2}}Y_{11} + \frac{\varepsilon_x + i\varepsilon_y}{\sqrt{2}}Y_{1-1} + \varepsilon_z Y_{10}\right)Y_{l_i m_i}(\theta,\varphi)d\Omega,\end{aligned} \tag{10.130}$$

where we have used $\langle \vec{r} \mid \psi_i \rangle = R_{n_i l_i}(\vec{r}) Y_{l_i m_i}(\Omega)$ and $\langle \vec{r} \mid \psi_f \rangle = R_{n_f l_f}(\vec{r}) Y_{l_f m_f}(\Omega)$.

The integration over the angular degrees of freedom can be calculated by means of the Wigner–Eckart theorem; we have shown in Chapter 7 that

$$\int d\Omega\, Y^*_{l_f m_f} Y_{1m'} Y_{l_i m_i} = \langle l_f,\ m_f | Y_{1m'} | l_i,\ m_i \rangle$$

$$= \sqrt{\frac{3(2l_i+1)}{4\pi(2l_f+1)}} \langle l_i, 1;\ 0, 0 | l_f,\ 0 \rangle \langle l_i, 1;\ m_i, m' | l_f,\ m_f \rangle. \tag{10.131}$$

Inserting (10.131) into (10.126) and (10.127), we obtain $\Gamma^{emi}_{i\to f} \sim \langle l_i, 1;\ m_i, m' \mid l_f,\ m_f \rangle^2$ and $\Gamma^{abs}_{i\to f} \sim \langle l_i, 1;\ m_i, m' \mid l_f,\ m_f \rangle^2$. Thus the *dipole selection rules* are specified by the selection rules of the Clebsch–Gordan coefficient $\langle l_i, 1;\ m_i, m' \mid l_f,\ m_f \rangle$:

- The transition rates are zero unless the values of m_f and m_i satisfy the condition $m_i + m' = m_f$ or $m'_f - m_i = m'$. But since m' takes only three values, $m' = -1, 0, 1$, we have

$$m_f - m_i = -1,\ 0,\ 1. \tag{10.132}$$

- The permissible values of l_f must lie between $l_i - 1$ and $l_i + 1$ (i.e., $l_i - 1 \le l_f \le l_i + 1$), so we have $-1 \le l_f - l_i \le 1$ or

$$l_f - l_i = -1,\ 0,\ 1. \tag{10.133}$$

  Note that, since the Clebsch–Gordan coefficient $\langle l_i, 1;\ m_i, m' \mid l_f,\ m_f \rangle$ vanishes for $l_i = l_f = 0$, no transition between $l_i = 0$ and $l_f = 0$ is allowed.

- Finally, since the coefficient $\langle l_i, 1;\ 0, 0 | l_f,\ 0 \rangle$ vanishes unless $(-1)^{l_i+1-l_f} = 1$ or $(-1)^{l_i-l_f} = -1$, then $(l_i - l_f)$ must be an odd integer:

$$l_f - l_i = \text{odd integer}. \tag{10.134}$$

  This means that, in the case of electric dipole transitions, the final and initial states must have different parities. As a result, electric dipole transitions like $1s \to 2s$, $2p \to 3p$, etc., are *forbidden*, while transitions like $1s \to 2p$, $2p \to 3s$, etc., are allowed.

10.5.6 Spontaneous Emission

It is clear from (10.126) that the rate of emission of a photon from an atom is not zero even in the absence of an external radiation field ($n_{\lambda,\vec{k}} = 0$). This corresponds to the *spontaneous emission of a photon*. The total transition rate corresponding to spontaneous emission can be inferred from (10.126) by taking $n_{\lambda,\vec{k}} = 0$:

$$\Gamma^{emi}_{i\to f} = \frac{4\pi^2\omega_{fi}^2}{\omega V} |\vec{\varepsilon}^*_\lambda \cdot \vec{d}_{fi}|^2 \delta(E_f - E_i + \hbar\omega), \tag{10.135}$$

where $\vec{d}_{fi}$ is the matrix element for the electron's electric dipole moment $\vec{d} = -e\vec{r}$:

$$\vec{d}_{fi} = \langle \psi_f \mid \vec{d} \mid \psi_i \rangle = -e \langle \psi_f \mid \vec{r} \mid \psi_i \rangle. \tag{10.136}$$

The relation (10.135) gives the transition probability per unit time corresponding to the transition of the atom from the initial state $|\psi_i\rangle$ to the final state $|\psi_f\rangle$ as a result of its spontaneous emission of a photon of energy $\hbar\omega$. Thus the final states of the system consist of products of discrete atomic states and a continuum of photonic states. The photon emitted will be detected in general as having a momentum in the momentum interval $(p, p+dp)$ located around $p = \hbar k = \hbar\omega/c$. The transition rate (10.135) needs then to be summed over the continuum of the final photonic states. The number of final photonic states within the unit volume V, whose momenta are within the interval $(p, p+dp)$, is given by

$$d^3n = \frac{Vd^3p}{(2\pi\hbar)^3} = \frac{Vp^2dp\,d\Omega}{(2\pi\hbar)^3} = \frac{V\hbar^3\omega^2}{(2\pi\hbar)^3c^3}d\Omega\,d\omega = \frac{V\omega^2}{(2\pi c)^3}d\Omega\,d\omega. \tag{10.137}$$

Thus, the transition rate corresponding to the emission of a photon in the solid angle $d\Omega$ is obtained by integrating (10.135) over $d\omega$:

$$\begin{aligned} dW^{emi}_{i\to f} &= \frac{V}{(2\pi)^3c^3}d\Omega\int\omega^2\Gamma^{emi}_{i\to f}d\omega = \frac{1}{2\pi c^3}|\vec{\varepsilon}^*_\lambda\cdot\vec{d}_{fi}|^2d\Omega\int\omega^2_{fi}\omega\delta(E_f-E_i+\hbar\omega)\,d\omega \\ &= \frac{1}{2\pi\hbar c^3}|\vec{\varepsilon}^*_\lambda\cdot\vec{d}_{fi}|^2d\Omega\int\omega^2_{fi}\omega\delta(\omega_{if}-\omega)\,d\omega, \end{aligned} \tag{10.138}$$

where we have used the fact $\delta(E_f - E_i + \hbar\omega) = (1/\hbar)\delta(\omega_{if} - \omega)$ with $\omega_{if} = (E_i - E_f)/\hbar$. Carrying out the integration, we can reduce (10.138) to

$$dW^{emi}_{i\to f} = \frac{\omega^3}{2\pi\hbar c^3}|\vec{\varepsilon}^*_\lambda\cdot\vec{d}_{fi}|^2d\Omega. \tag{10.139}$$

The transition rate (10.139) corresponds to a specific polarization; that is, the photon emitted travels along the direction $\vec{n}$ (since $\vec{k} = k\vec{n}$), which is normal to $\vec{\varepsilon}^*_\lambda$. To find the transition rate corresponding to any polarization, we need to sum over the two polarizations of the photon:

$$\sum_{\lambda=1}^{2}|\vec{\varepsilon}^*_\lambda\cdot\vec{d}_{fi}|^2 = |\varepsilon^*_1(d_{fi})_1|^2 + |\varepsilon^*_2(d_{fi})_2|^2 = |\vec{d}_{fi}|^2 - |(d_{fi})_3|^2. \tag{10.140}$$

Since the three directions of $\vec{d}_{fi}$ are equivalent, we have

$$\langle|(d_{fi})_1|^2\rangle = \langle|(d_{fi})_2|^2\rangle = \langle|(d_{fi})_3|^2\rangle = \frac{1}{3}\langle|\vec{d}_{fi}|^2\rangle. \tag{10.141}$$

Thus, an average over polarization yields

$$\sum_{\lambda=1}^{2}|\vec{\varepsilon}^*_\lambda\cdot\vec{d}_{fi}|^2 = |\vec{d}_{fi}|^2 - \frac{1}{3}|\vec{d}_{fi}|^2 = \frac{2}{3}|\vec{d}_{fi}|^2. \tag{10.142}$$

Substituting (10.142) into (10.139), we obtain the average transition rate corresponding to the emission of the photon into the solid angle $d\Omega$:

$$dW^{emi}_{i\to f} = \frac{\omega^3}{3\pi\hbar c^3}|\vec{d}_{fi}|^2d\Omega. \tag{10.143}$$

An integration over all possible (photonic) directions ($|\vec{d}_{fi}|^2$ is not included in the integration since we are integrating over the angular part of the photonic degrees of freedom only and not over the electron's) yields $\int d\Omega = 4\pi$. Thus, the transition rate associated with the emission of the photon is

$$\boxed{W_{i\to f}^{emi} = \frac{4}{3}\frac{\omega^3}{\hbar c^3}\left|\vec{d}_{fi}\right|^2 = \frac{4}{3}\frac{\omega^3 e^2}{\hbar c^3}\left|\langle\psi_f \mid \vec{r}\mid\psi_i\rangle\right|^2,} \tag{10.144}$$

where $\omega = (E_f - E_i)/\hbar$.

The total *power (or intensity) radiated* by the electron is obtained by multiplying the total rates (10.144) by $\hbar\omega$:

$$\boxed{I_{i\to f} = \hbar\omega W_{i\to f}^{emi} = \frac{4}{3}\frac{\omega^4}{c^3}|\vec{d}_{fi}|^2 = \frac{4}{3}\frac{\omega^4 e^2}{c^3}\left|\langle\psi_f \mid \vec{r}\mid\psi_i\rangle\right|^2.} \tag{10.145}$$

The transition rates derived above, (10.144) and (10.145), were obtained for single-electron atoms. For atoms that have Z electrons, we must replace the dipole moment $\vec{d} = -e\vec{r}$ with the dipole moment of all Z electrons: $\vec{d} = -e\sum_{j=1}^{Z}\vec{r}_j$.

The mean lifetime τ of an excited state can be obtained by adding together the total transition probabilities per unit time (10.144) for all possible final states:

$$\boxed{\tau = \frac{1}{W} = \frac{1}{\sum_f W_{i\to f}}.} \tag{10.146}$$

Example 10.3

A particle of charge q and mass m is moving in a one-dimensional harmonic oscillator potential of frequency ω_0.

(a) Find the rate of spontaneous emission for a transition from an excited state $\mid n\rangle$ to the ground state.

(b) Obtain an estimate for the rate calculated in (a) and the lifetime of the state $\mid n\rangle$ when the particle is an electron and $\omega_0 = 3\times 10^{14}\,\mathrm{rad\,s^{-1}}$.

(c) Find the condition under which the dipole approximation is valid for the particle of (b).

Solution

(a) The spontaneous emission rate for a transition from an excited state $\mid n\rangle$ to $\mid 0\rangle$ is given by (10.144):

$$W_{n\to 0}^{emi} = \frac{4}{3}\frac{\omega^3 q^2}{\hbar c^3}\left|\langle 0 \mid \hat{X} \mid n\rangle\right|^2, \tag{10.147}$$

where $\omega = (E_n - E_0)/\hbar = (n+\frac{1}{2})\omega_0 - \frac{1}{2}\omega_0 = n\omega_0$. Since $\hat{a}\mid n\rangle = \sqrt{n}\mid n-1\rangle$ and $\hat{a}^\dagger\mid n\rangle = \sqrt{n+1}\mid n+1\rangle$, and since $\hat{X} = \sqrt{\hbar/(2m\omega_0)}(\hat{a}^\dagger + \hat{a})$, we have

$$\langle 0 \mid \hat{X} \mid n\rangle = \sqrt{\frac{\hbar}{2m\omega_0}}\langle 0 \mid \hat{a}^\dagger + \hat{a} \mid n\rangle = \sqrt{\frac{\hbar}{2m\omega_0}}\left[\sqrt{n+1}\delta_{0,n+1} + \sqrt{n}\delta_{0,n-1}\right]. \tag{10.148}$$

Thus only a transition from $|\,1\rangle$ to $|\,0\rangle$ is possible; hence $n = 1$, $\omega = \omega_0$, and $\langle 0 \mid \hat{X} \mid 1\rangle = \sqrt{\hbar/(2m\omega_0)}$. The emission rate (10.147) then becomes

$$W_{1\to 0}^{emi} = \frac{4}{3}\frac{\omega^3 q^2}{\hbar c^3}\left|\langle 0 \mid \hat{X} \mid 1\rangle\right|^2 = \frac{4}{3}\frac{\omega_0^3 q^2}{\hbar c^3}\frac{\hbar}{2m\omega_0} = \frac{2}{3}\frac{\omega_0^2 q^2}{mc^3}. \tag{10.149}$$

(b) If the particle is an electron, we have $q = -e$:

$$W_{1\to 0}^{emi} = \frac{2}{3}\frac{\omega_0^2 e^2}{m_e c^3} = \frac{2\alpha}{3}\frac{\omega_0^2 \hbar}{m_e c^2} = \frac{2\alpha}{3}\frac{\hbar c}{m_e c^2}\frac{\omega_0^2}{c}. \tag{10.150}$$

Using $m_e c^2 = 0.511\,\text{MeV}$, $\hbar c = 197.33\,\text{MeV fm}$, we have

$$W_{1\to 0}^{emi} = \frac{2\alpha}{3}\frac{\hbar c}{m_e c^2}\frac{\omega_0^2}{c} = \frac{2}{3\times 137}\frac{197.33\,\text{MeV fm}}{0.511\,\text{MeV}}\frac{9\times 10^{28}\,\text{s}^{-2}}{3\times 10^8\,\text{m s}^{-1}} = 5.6\times 10^5\,\text{s}^{-1}. \tag{10.151}$$

The lifetime of the $|\,1\rangle$ state is

$$\tau = \frac{1}{W_{1\to 0}^{emi}} = \frac{3}{2}\frac{m_e c^3}{\omega_0^2 e^2} = \frac{1}{5.6\times 10^5 sec^{-2}} = 0.18\times 10^{-5}\,\text{s}. \tag{10.152}$$

(c) For the dipole approximation to be valid, we need $kx \ll 1$, where x was calculated in (10.148) for $n = 1$: $x = \sqrt{\hbar/(2m_e\omega_0)}$. As for k, a crude estimate yields $k = \omega/c = (E_1 - E_0)/(\hbar c) = \omega_0/c$. Thus, we have

$$kx = \frac{\omega_0}{c}\sqrt{\frac{\hbar}{2m_e\omega_0}} = \sqrt{\frac{\hbar\omega_0}{2m_e c^2}} \ll 1 \qquad \Longrightarrow \qquad \hbar\omega_0 \ll 2m_e c^2. \tag{10.153}$$

This is indeed the case since $2m_e c^2 = 1.022\,\text{MeV}$ is very large compared to

$$\hbar\omega_0 = \hbar c\frac{\omega_0}{c} = 197.33\,\text{MeV fm}\times\frac{3\times 10^{14}\,\text{s}^{-1}}{3\times 10^8\,\text{m s}^{-1}} = 2.0\times 10^{-7}\,\text{MeV}. \tag{10.154}$$

10.6 Solved Problems

Problem 10.1

(a) Calculate the position and the momentum operators, $\hat{X}_H(t)$ and $\hat{P}_H(t)$, in the Heisenberg picture for a one-dimensional harmonic oscillator.

(b) Find the Heisenberg equations of motion for $\hat{X}_H(t)$ and $\hat{P}_H(t)$.

Solution

In the Schrödinger picture, where the operators do not depend explicitly on time, the Hamiltonian of a one-dimensional harmonic oscillator is given by

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{X}^2. \tag{10.155}$$

(a) Using the commutation relations

$$[\hat{H},\ \hat{X}] = \frac{1}{2m}[\hat{P}^2,\ \hat{X}] = -\frac{i\hbar}{m}\hat{P}, \tag{10.156}$$

$$[\hat{H},\ \hat{P}] = \frac{1}{2}m\omega^2[\hat{X}^2,\ \hat{P}] = i\hbar m\omega^2\hat{X}, \tag{10.157}$$

along with

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A},\ \hat{B}] + \frac{1}{2!}[\hat{A},\ [\hat{A},\ \hat{B}]] + \frac{1}{3!}[\hat{A},\ [\hat{A},\ [\hat{A},\ \hat{B}]]] + \cdots, \tag{10.158}$$

we may write (see Eq. (10.11))

$$\begin{aligned}\hat{X}_H(t) &= e^{it\hat{H}/\hbar}\hat{X}e^{-it\hat{H}/\hbar} = \hat{X} + \frac{it}{\hbar}[\hat{H},\ \hat{X}] + \frac{1}{2!}\left(\frac{it}{\hbar}\right)^2[\hat{H},\ [\hat{H},\ \hat{X}]] + \cdots \\ &= \hat{X} + \frac{t}{m}\hat{P} - \frac{(\omega t)^2}{2!}\hat{X} - \frac{(\omega t)^3}{3!}\frac{1}{m\omega}\hat{P} + \frac{(\omega t)^4}{4!}\hat{X} + \frac{(\omega t)^5}{5!}\frac{1}{m\omega}\hat{P} + \cdots \\ &= \hat{X}\left[1 - \frac{(\omega t)^2}{2!} + \frac{(\omega t)^4}{4!} + \cdots\right] + \frac{1}{m\omega}\hat{P}\left[(\omega t) - \frac{(\omega t)^3}{3!} + \frac{(\omega t)^5}{5!} + \cdots\right],\end{aligned} \tag{10.159}$$

or

$$\boxed{\hat{X}_H(t) = \hat{X}\cos(\omega t) + \frac{1}{m\omega}\hat{P}\sin(\omega t).} \tag{10.160}$$

A similar calculation yields (see Eq. (10.11))

$$\begin{aligned}\hat{P}_H(t) &= e^{it\hat{H}/\hbar}\hat{P}e^{-it\hat{H}/\hbar} = \hat{P} + \frac{it}{\hbar}[\hat{H},\ \hat{P}] + \frac{1}{2!}\left(\frac{it}{\hbar}\right)^2[\hat{H},\ [\hat{H},\ \hat{P}]] + \cdots \\ &= \hat{P}\left[1 - \frac{(\omega t)^2}{2!} + \frac{(\omega t)^4}{4!} + \cdots\right] - m\omega\hat{X}\left[(\omega t) - \frac{(\omega t)^3}{3!} + \frac{(\omega t)^5}{5!} + \cdots\right],\end{aligned} \tag{10.161}$$

or

$$\boxed{\hat{P}_H(t) = \hat{P}\cos(\omega t) - m\omega\hat{X}\sin(\omega t).} \tag{10.162}$$

(b) To find the equations of motion of $\hat{X}_H(t)$ and $\hat{P}_H(t)$, we need to use the Heisenberg equation $d\hat{A}_H(t)/dt = (1/i\hbar)[\hat{A}_H(t),\ \hat{H}]$ which, along with (10.156) and (10.157), leads to

$$\frac{d\hat{X}_H(t)}{dt} = \frac{1}{i\hbar}[\hat{X}_H(t),\ \hat{H}] = \frac{1}{i\hbar}e^{it\hat{H}/\hbar}[\hat{X},\ \hat{H}]e^{-it\hat{H}/\hbar} = \frac{1}{i\hbar}\frac{i\hbar}{m}e^{it\hat{H}/\hbar}\hat{P}e^{-it\hat{H}/\hbar}, \tag{10.163}$$

$$\frac{d\hat{P}_H(t)}{dt} = \frac{1}{i\hbar}[\hat{P}_H(t),\ \hat{H}] = \frac{1}{i\hbar}e^{it\hat{H}/\hbar}[\hat{P},\ \hat{H}]e^{-it\hat{H}/\hbar} = \frac{(-i\hbar m\omega^2)}{i\hbar}e^{it\hat{H}/\hbar}\hat{X}e^{-it\hat{H}/\hbar}, \tag{10.164}$$

or

$$\boxed{\frac{d\hat{X}_H(t)}{dt} = \frac{1}{m}\hat{P}_H(t), \qquad \frac{d\hat{P}_H(t)}{dt} = -m\omega^2\hat{X}_H(t).} \tag{10.165}$$

Problem 10.2
Using the expressions derived in Problem 10.1 for $\hat{X}_H(t)$ and $\hat{P}_H(t)$, evaluate the following commutators for a harmonic oscillator:

$$\left[\hat{X}_H(t_1), \hat{P}_H(t_2)\right], \quad \left[\hat{X}_H(t_1), \hat{X}_H(t_2)\right], \quad \left[\hat{P}_H(t_1), \hat{P}_H(t_2)\right].$$

Solution
Using (10.160) and (10.162) along with the commutation relations $[\hat{X}, \hat{P}] = i\hbar$ and $[\hat{X}, \hat{X}] = [\hat{P}, \hat{P}] = 0$, we have

$$\begin{aligned}
[\hat{X}_H(t_1), \hat{P}_H(t_2)] &= \left[\hat{X}\cos(\omega t_1) + \frac{1}{m\omega}\hat{P}\sin(\omega t_1), \hat{P}\cos(\omega t_2) - m\omega\hat{X}\sin(\omega t_2)\right] \\
&= [\hat{X}, \hat{P}]\cos(\omega t_1)\cos(\omega t_2) - [\hat{P}, \hat{X}]\sin(\omega t_1)\sin(\omega t_2) \\
&= i\hbar\left[\cos(\omega t_1)\cos(\omega t_2) + \sin(\omega t_1)\sin(\omega t_2)\right], \qquad (10.166)
\end{aligned}$$

or

$$\boxed{[\hat{X}_H(t_1), \hat{P}_H(t_2)] = i\hbar\cos\left[\omega(t_1 - t_2)\right].} \qquad (10.167)$$

A similar calculation yields

$$\begin{aligned}
[\hat{X}_H(t_1), \hat{X}_H(t_2)] &= \left[\hat{X}\cos(\omega t_1) + \frac{1}{m\omega}\hat{P}\sin(\omega t_1), \hat{X}\cos(\omega t_2) + \frac{1}{m\omega}\hat{P}\sin(\omega t_2)\right] \\
&= \frac{1}{m\omega}[\hat{X}, \hat{P}]\cos(\omega t_1)\sin(\omega t_2) + \frac{1}{m\omega}[\hat{P}, \hat{X}]\sin(\omega t_1)\cos(\omega t_2) \\
&= \frac{i\hbar}{m\omega}\left[\cos(\omega t_1)\sin(\omega t_2) - \sin(\omega t_1)\cos(\omega t_2)\right], \qquad (10.168)
\end{aligned}$$

or

$$\boxed{[\hat{X}_H(t_1), \hat{X}_H(t_2)] = -\frac{i\hbar}{m\omega}\sin\left[\omega(t_1 - t_2)\right].} \qquad (10.169)$$

Similarly, we have

$$\begin{aligned}
[\hat{P}_H(t_1), \hat{P}_H(t_2)] &= \left[\hat{P}\cos(\omega t_1) - m\omega\hat{X}\sin(\omega t_1), \hat{P}\cos(\omega t_2) - m\omega\hat{X}\sin(\omega t_2)\right] \\
&= -m\omega[\hat{P}, \hat{X}]\cos(\omega t_1)\sin(\omega t_2) - m\omega[\hat{X}, \hat{P}]\sin(\omega t_1)\cos(\omega t_2) \\
&= -i\hbar m\omega\left[\sin(\omega t_1)\cos(\omega t_2) - \cos(\omega t_1)\sin(\omega t_2)\right], \qquad (10.170)
\end{aligned}$$

or

$$\boxed{[\hat{P}_H(t_1), \hat{P}_H(t_2)] = -i\hbar m\omega\sin\left[\omega(t_1 - t_2)\right].} \qquad (10.171)$$

Problem 10.3
Evaluate the quantity $\langle n \mid \hat{X}_H(t)\hat{X} \mid n\rangle$ for the nth excited state of a one-dimensional harmonic oscillator, where $\hat{X}_H(t)$ and $\hat{X}$ designate the position operators in the Heisenberg picture and the Schrödinger picture.

Solution

Using the expression of $\hat{X}_H(t)$ calculated in (10.160), we have

$$\langle n \mid \hat{X}_H(t)\hat{X} \mid n\rangle = \langle n \mid \hat{X}^2 \mid n\rangle \cos(\omega t) + \frac{1}{m\omega}\langle n \mid \hat{P}\hat{X} \mid n\rangle \sin(\omega t). \tag{10.172}$$

Since, for a harmonic oscillator, $\hat{X}$ and $\hat{P}$ are given by

$$\hat{X} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}^\dagger + \hat{a}), \qquad \hat{P} = i\sqrt{\frac{m\hbar\omega}{2}}(\hat{a}^\dagger - \hat{a}), \tag{10.173}$$

and $\hat{a}^\dagger \mid n\rangle = \sqrt{n+1} \mid n+1\rangle$ and $\hat{a} \mid n\rangle = \sqrt{n} \mid n-1\rangle$, we have

$$\langle n \mid \hat{X}^2 \mid n\rangle = \frac{\hbar}{2m\omega}\langle n \mid \hat{a}^{\dagger\,2} + \hat{a}^2 + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} \mid n\rangle = \frac{\hbar}{2m\omega}(2n+1), \tag{10.174}$$

$$\langle n \mid \hat{P}\hat{X} \mid n\rangle = \frac{i\hbar}{2}\langle n \mid \hat{a}^{\dagger\,2} + \hat{a}^2 - \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} \mid n\rangle = -\frac{i\hbar}{2}, \tag{10.175}$$

since $\langle n \mid \hat{a}^{\dagger\,2} \mid n\rangle = \langle n \mid \hat{a}^2 \mid n\rangle = 0$, $\langle n \mid \hat{a}^\dagger\hat{a} \mid n\rangle = n$ and $\langle n \mid \hat{a}\hat{a}^\dagger \mid n\rangle = n+1$. Inserting (10.174) and (10.175) into (10.172), we obtain

$$\langle n \mid \hat{X}_H(t)\hat{X} \mid n\rangle = \frac{\hbar}{2m\omega}\Big[(2n+1)\cos(\omega t) - i\sin(\omega t)\Big]. \tag{10.176}$$

Problem 10.4
The Hamiltonian due to the interaction of a particle of mass m, charge q, and spin $\vec{S}$ with a magnetic field pointing along the z-axis is $\hat{H} = -(qB/mc)\hat{S}_z$. Write the Heisenberg equations of motion for the time-dependent spin operators $\hat{S}_x(t)$, $\hat{S}_y(t)$, and $\hat{S}_z(t)$, and solve them to obtain the operators as functions of time.

Solution
Let us write $\hat{H}$ in a lighter form $\hat{H} = \omega\hat{S}_z$ where $\omega = -qB/mc$. The commutation of $\hat{H}$ with the components of the spin operator can be inferred at once from $[\hat{S}_x, \hat{S}_z] = -i\hbar\hat{S}_y$ and $[\hat{S}_y, \hat{S}_z] = i\hbar\hat{S}_x$:

$$[\hat{S}_x, \hat{H}] = -i\hbar\omega\hat{S}_y, \qquad [\hat{S}_y, \hat{H}] = i\hbar\omega\hat{S}_x, \qquad [\hat{S}_z, \hat{H}] = 0. \tag{10.177}$$

The Heisenberg equations of motion for $\hat{S}_x(t)$, $\hat{S}_y(t)$, and $\hat{S}_z(t)$ can be obtained from $d\hat{A}_H(t)/dt = (1/i\hbar)[\hat{A}_H(t), \hat{H}] = (1/i\hbar)e^{it\hat{H}/\hbar}[\hat{A}(0), \hat{H}]e^{-it\hat{H}/\hbar}$ which, using (10.177), leads to

$$\begin{aligned}\frac{d\hat{S}_x(t)}{dt} &= \frac{1}{i\hbar}[\hat{S}_x(t), \hat{H}] = \frac{1}{i\hbar}e^{it\hat{H}/\hbar}[\hat{S}_x(0), \hat{H}]e^{-it\hat{H}/\hbar} \\ &= \frac{-i\hbar\omega}{i\hbar}e^{it\hat{H}/\hbar}\hat{S}_y(0)e^{-it\hat{H}/\hbar} = -\omega\hat{S}_y(t).\end{aligned} \tag{10.178}$$

Similarly, we have

$$\frac{d\hat{S}_y(t)}{dt} = \frac{1}{i\hbar}e^{it\hat{H}/\hbar}[\hat{S}_y(0), \hat{H}]e^{-it\hat{H}/\hbar} = \frac{i\hbar\omega}{i\hbar}e^{it\hat{H}/\hbar}\hat{S}_x(0)e^{-it\hat{H}/\hbar} = \omega\hat{S}_x(t), \tag{10.179}$$

$$\frac{d\hat{S}_z(t)}{dt} = \frac{1}{i\hbar}e^{it\hat{H}/\hbar}[\hat{S}_z(0), \hat{H}]e^{-it\hat{H}/\hbar} = 0. \tag{10.180}$$

To solve (10.178) and (10.179), we may combine them into two more conducive equations:

$$\frac{d\hat{S}_{\pm}(t)}{dt} = \pm i\omega \hat{S}_{\pm}(t), \tag{10.181}$$

where $\hat{S}_{\pm}(t) = \hat{S}_x(t) \pm i\hat{S}_y(t)$. The solutions of (10.181) are $\hat{S}_{\pm}(t) = \hat{S}_{\pm}(0)e^{\pm i\omega t}$ which, when combined with $\hat{S}_x(t) = \frac{1}{2}[\hat{S}_+(t) + \hat{S}_-(t)]$ and $\hat{S}_y(t) = \frac{1}{2i}[\hat{S}_+(t) - \hat{S}_-(t)]$, lead to

$$\hat{S}_x(t) = \hat{S}_x(0)\cos(\omega t) - \hat{S}_y(0)\sin(\omega t), \tag{10.182}$$

$$\hat{S}_y(t) = \hat{S}_y(0)\cos(\omega t) + \hat{S}_x(0)\sin(\omega t). \tag{10.183}$$

The solution of (10.180) is obvious:

$$\frac{d\hat{S}_z(t)}{dt} = 0 \quad \Longrightarrow \quad \hat{S}_z(t) = \hat{S}_z(0). \tag{10.184}$$

Problem 10.5

Consider a spinless particle of mass m, which is moving in a one-dimensional infinite potential well with walls at $x = 0$ and $x = a$.

(a) Find $\hat{X}_H(t)$ and $\hat{P}_H(t)$ in the Heisenberg picture.

(b) If at $t = 0$ the particle is in the state $\psi(x, 0) = [\phi_1(x) + \phi_2(x)]/\sqrt{2}$, where $\phi_1(x)$ and $\phi_2(x)$ are the ground and first excited states, respectively, with $\phi_n(x) = \sqrt{2/a}\sin(n\pi x/a)$, find the state vector $\psi(x, t)$ for $t > 0$ in the Schrödinger picture.

(c) Evaluate $\langle\psi(x,t) \mid \hat{X} \mid \psi(x,t)\rangle$ and $\langle\psi(x,t) \mid \hat{P} \mid \psi(x,t)\rangle$ as a function of time in the Schrödinger picture.

(d) Evaluate $\langle\psi(x,t) \mid \hat{X}_H(t) \mid \psi(x,t)\rangle$ and $\langle\psi(x,t) \mid \hat{P}_H(t) \mid \psi(x,t)\rangle$ as a function of time in the Schrödinger picture.

Solution

(a) Since the particle's Hamiltonian is purely kinetic, $\hat{H} = \hat{P}^2/2m$, we have $[\hat{H}, \hat{P}] = 0$ and

$$[\hat{H}, \hat{X}] = \frac{1}{2m}[\hat{P}^2, \hat{X}] = -\frac{i\hbar}{m}\hat{P}. \tag{10.185}$$

Using these relations along with (10.158), we obtain

$$\hat{X}_H(t) = e^{it\hat{H}/\hbar}\hat{X}e^{-it\hat{H}/\hbar} = \hat{X} + \frac{it}{\hbar}[\hat{H}, \hat{X}] + \frac{1}{2!}\left(\frac{it}{\hbar}\right)^2[\hat{H}, [\hat{H}, \hat{X}]] + \cdots, \tag{10.186}$$

and since $[\hat{H}, [\hat{H}, \hat{X}]] = -(i\hbar/m)[\hat{H}, \hat{P}] = 0$, we end up with

$$\hat{X}_H(t) = \hat{X} + \frac{t}{m}\hat{P}. \tag{10.187}$$

On the other hand, since $[\hat{H}, \hat{P}] = 0$, we have

$$\hat{P} = \hat{P}_H(t). \tag{10.188}$$

(b) Since the energy of the nth level is given by $E_n = n^2\pi^2\hbar^2/(2ma^2)$, we have

$$\begin{aligned}\psi(x,t) &= \frac{1}{\sqrt{2}}\left[\phi_1(x)e^{-iE_1t/\hbar} + \phi_2(x)e^{-iE_2t/\hbar}\right] \\ &= \frac{1}{\sqrt{a}}\left[e^{-iE_1t/\hbar}\sin\left(\frac{\pi x}{a}\right) + e^{-iE_2t/\hbar}\sin\left(\frac{2\pi x}{a}\right)\right].\end{aligned} \tag{10.189}$$

(c) Using (10.189) we can write

$$\begin{aligned}\langle\psi(x,t) \mid \hat{X} \mid \psi(x,t)\rangle = \frac{1}{2}\big[&\langle\phi_1 \mid \hat{X} \mid \phi_1\rangle + \langle\phi_2 \mid \hat{X} \mid \phi_2\rangle + \langle\phi_1 \mid \hat{X} \mid \phi_2\rangle e^{-i(E_2-E_1)t/\hbar} \\ &+ \langle\phi_2 \mid \hat{X} \mid \phi_1\rangle e^{i(E_2-E_1)t/\hbar}\big].\end{aligned} \tag{10.190}$$

Since $\langle\phi_n|\hat{X} \mid \phi_n\rangle = a/2$ (Chapter 4) and

$$\langle\phi_1 \mid \hat{X} \mid \phi_2\rangle = \langle\phi_2 \mid \hat{X} \mid \phi_1\rangle = \frac{2}{a}\int_0^a x\sin\left(\frac{\pi x}{a}\right)\sin\left(\frac{2\pi x}{a}\right)dx = -\frac{16a}{9\pi^2}, \tag{10.191}$$

we can rewrite (10.190) as

$$\begin{aligned}\langle\psi(x,t) \mid \hat{X} \mid \psi(x,t)\rangle &= \frac{1}{2}\left[\frac{a}{2} + \frac{a}{2} - \frac{16a}{9\pi^2}\left(e^{-i(E_2-E_1)t/\hbar} + e^{i(E_2-E_1)t/\hbar}\right)\right] \\ &= \frac{a}{2} - \frac{16a}{9\pi^2}\cos\left(\frac{3\pi^2\hbar t}{2ma^2}\right),\end{aligned} \tag{10.192}$$

since $E_2 - E_1 = 3\pi^2\hbar^2/(2ma^2)$.

A similar calculation which uses $\langle\phi_n|\hat{P} \mid \phi_n\rangle = 0$ and

$$\langle\phi_1 \mid \hat{P} \mid \phi_2\rangle = -i\hbar\frac{4\pi}{a^2}\int_0^a \sin\left(\frac{\pi x}{a}\right)\cos\left(\frac{2\pi x}{a}\right)dx = \frac{8i\hbar}{3a} = -\langle\phi_2 \mid \hat{P} \mid \phi_1\rangle \tag{10.193}$$

leads to

$$\begin{aligned}\langle\psi(x,t) \mid \hat{P} \mid \psi(x,t)\rangle = \frac{1}{2}\big[&\langle\phi_1 \mid \hat{P} \mid \phi_1\rangle + \langle\phi_2 \mid \hat{P} \mid \phi_2\rangle + \langle\phi_1 \mid \hat{P} \mid \phi_2\rangle e^{-i(E_2-E_1)t/\hbar} \\ &+ \langle\phi_2 \mid \hat{P} \mid \phi_1\rangle e^{i(E_2-E_1)t/\hbar}\big],\end{aligned} \tag{10.194}$$

or to

$$\langle\psi(x,t) \mid \hat{P} \mid \psi(x,t)\rangle = \frac{1}{2}\left[\frac{8i\hbar}{3a}e^{-i(E_2-E_1)t/\hbar} - \frac{8i\hbar}{3a}e^{i(E_2-E_1)t/\hbar}\right] = \frac{8\hbar}{3a}\sin\left(\frac{3\pi^2\hbar t}{2ma^2}\right). \tag{10.195}$$

(d) From (10.187) we have

$$\langle\psi(x,t) \mid \hat{X}_H(t) \mid \psi(x,t)\rangle = \langle\psi(x,t) \mid \hat{X} \mid \psi(x,t)\rangle + \frac{t}{m}\langle\psi(x,t) \mid \hat{P} \mid \psi(x,t)\rangle. \tag{10.196}$$

Inserting the expressions for $\langle\psi(x,t) \mid \hat{X} \mid \psi(x,t)\rangle$ and $\langle\psi(x,t) \mid \hat{P} \mid \psi(x,t)\rangle$ calculated in (10.192) and (10.195), we obtain

$$\langle\psi(x,t) \mid \hat{X}_H(t) \mid \psi(x,t)\rangle = \frac{a}{2} - \frac{16a}{9\pi^2}\cos\left(\frac{3\pi^2\hbar t}{2ma^2}\right) + \frac{8\hbar t}{3ma}\sin\left(\frac{3\pi^2\hbar t}{2ma^2}\right), \tag{10.197}$$

and $\langle\psi(x,t) \mid \hat{P}_H(t) \mid \psi(x,t)\rangle$ is given by (10.195):

$$\langle\psi(x,t) \mid \hat{P}_H(t) \mid \psi(x,t)\rangle = \langle\psi(x,t) \mid \hat{P} \mid \psi(x,t)\rangle = \frac{8\hbar}{3a}\sin\left(\frac{3\pi^2\hbar t}{2ma^2}\right), \tag{10.198}$$

since, as shown in (10.188), we have $\hat{P}_H(t) = \hat{P}$.

Problem 10.6
A particle, initially (i.e., $t \to -\infty$) in its ground state in an infinite potential well whose walls are located at $x = 0$ and $x = a$, is subject at time $t = 0$ to a time-dependent perturbation $\hat{V}(t) = \varepsilon\hat{x}e^{-t^2}$ where ε is a small real number. Calculate the probability that the particle will be found in its first excited state after a sufficiently long time (i.e., $t \to \infty$).

Solution
The transition probability from the ground state $n = 1$ (where $t \to -\infty$) to the first excited state $n = 2$ (where $t \to \infty$) is given by (10.43):

$$P_{1\to 2} = \frac{1}{\hbar^2}\left|\int_{-\infty}^{+\infty}\langle\psi_2|\hat{V}(t)|\psi_1\rangle e^{i\omega_{21}t}dt\right|^2, \tag{10.199}$$

where

$$\omega_{21} = \frac{E_2 - E_1}{\hbar} = \frac{4\pi^2\hbar}{2ma^2} - \frac{\pi^2\hbar}{2ma^2} = \frac{3\pi^2\hbar}{2ma^2}, \tag{10.200}$$

$$\langle\psi_2|\hat{V}(t) \mid \psi_1\rangle = \frac{2\varepsilon}{a}e^{-t^2}\int_0^a x\sin\left(\frac{2\pi x}{a}\right)\sin\left(\frac{\pi x}{a}\right)dx = -\frac{16\varepsilon a}{9\pi^2}e^{-t^2}, \tag{10.201}$$

since $E_n = n^2\pi^2\hbar^2/(2ma^2)$ and $\psi_n(x) = \sqrt{2/a}\sin(n\pi x/a)$. Inserting (10.200) and (10.201) into (10.199), we have

$$P_{1\to 2} = \left(\frac{16\varepsilon a}{9\pi^2\hbar}\right)^2\left|\int_{-\infty}^{+\infty}e^{i\omega_{21}t-t^2}dt\right|^2. \tag{10.202}$$

A variable change $y = t - \frac{i}{2}\omega_{21}$ yields $i\omega_{21}t - t^2 = -\omega_{21}^2/4 - y^2$ and $dt = dy$:

$$P_{1\to 2} = \left(\frac{16\varepsilon a}{9\pi^2\hbar}\right)^2\left|e^{-\omega_{21}^2/4}\int_{-\infty}^{+\infty}e^{-y^2}dy\right|^2 = \pi\left(\frac{16\varepsilon a}{9\pi^2\hbar}\right)^2\exp\left(-\frac{9\pi^4\hbar^2}{8m^2a^4}\right), \tag{10.203}$$

since $\omega_{21} = 3\pi^2\hbar/(2ma^2)$.

Problem 10.7
A particle is initially (i.e., $t = 0$) in its ground state in a one-dimensional harmonic oscillator potential. At $t = 0$ a perturbation $\hat{V}(x,t) = V_0\hat{x}^3e^{-t/\tau}$ is turned on. Calculate to first order the probability that, after a sufficiently long time (i.e., $t \to \infty$), the system will have made a transition to a given excited state; consider all final states.

Solution

The transition probability from the ground state $n = 0$ to an excited state n is given by (10.43):

$$P_{0\to n} = \frac{1}{\hbar^2}\left|\int_0^{+\infty} \langle n \mid \hat{V}(t)|0\rangle e^{i\omega_{n0}t}dt\right|^2 = \frac{V_0^2}{\hbar^2}\left|\langle n \mid \hat{x}^3|0\rangle\right|^2\left|\int_0^{+\infty} e^{-(1/\tau - in\omega)t}dt\right|^2, \tag{10.204}$$

where $\omega_{n0} = \frac{E_n - E_0}{\hbar} = n\omega$ (since $E_n = \hbar\omega(n + \frac{1}{2})$) and the time integration was calculated in (10.65):

$$\left|\int_0^{\infty} e^{-(1/\tau - in\omega)t}dt\right|^2 = \frac{1}{n^2\omega^2 + 1/\tau^2}. \tag{10.205}$$

Since $\hat{a} \mid n\rangle = \sqrt{n} \mid n-1\rangle$ and $\hat{a}^\dagger \mid n\rangle = \sqrt{n+1} \mid n+1\rangle$, and since $\hat{X}^3 = (\hbar/2m\omega)^{3/2}(\hat{a}^\dagger + \hat{a})(\hat{a}^2 + \hat{a}^{\dagger 2} + 2\hat{a}^\dagger\hat{a} + 1)$, the only terms that survive in $\langle n \mid \hat{x}^3|0\rangle$ are

$$\langle n \mid \hat{X}^3 \mid 0\rangle = \left(\frac{\hbar}{2m\omega}\right)^{3/2}\langle n \mid \hat{a}^{\dagger 3} + \hat{a}\hat{a}^{\dagger 2} + \hat{a}^\dagger \mid 0\rangle = \left(\frac{\hbar}{2m\omega}\right)^{3/2}\left(\sqrt{6}\delta_{n,3} + 3\delta_{n,1}\right). \tag{10.206}$$

This implies that the particle can be found after a long duration only either in the first or in the third excited state.

Inserting (10.205) and (10.206) into (10.204), we can verify that the probabilities corresponding to the transitions from the ground state to the first, the second and the third excited states are given, respectively, by

$$P_{0\to 1} = \frac{V_0^2}{\hbar^2}\left|\langle 1 \mid \hat{x}^3|0\rangle\right|^2\left|\int_0^{+\infty} e^{-(1/\tau - i\omega)t}dt\right|^2 = \left(\frac{\hbar}{2m\omega}\right)^3 \frac{9V_0^2}{(\hbar\omega)^2 + \hbar^2/\tau^2}, \tag{10.207}$$

$$P_{0\to 2} = 0 \tag{10.208}$$

$$P_{0\to 3} = \frac{V_0^2}{\hbar^2}\left|\langle 3|\hat{x}^3|0\rangle\right|^2\left|\int_0^{+\infty} e^{-(1/\tau - 3i\omega)t}dt\right|^2 = \left(\frac{\hbar}{2m\omega}\right)^3 \frac{6V_0^2}{(3\hbar\omega)^2 + \hbar^2/\tau^2}. \tag{10.209}$$

Therefore the system cannot undergo transitions to the second excited state nor to excited states higher than $n = 3$; that is, $P_{0\to 2} = 0$, since $\langle 2 \mid \hat{X}^3 \mid 0\rangle = 0$ and $P_{0\to n} = 0$ when $n > 3$, since $\langle n \mid \hat{X}^3 \mid 0\rangle = 0$ for $n > 3$.

Problem 10.8

A hydrogen atom, initially (i.e., $t \to -\infty$) in its ground state, is placed starting at time $t = 0$ in a time-dependent electric field pointing along the z-axis $\vec{E}(t) = E_0\tau\vec{k}/(\tau^2 + t^2)$, where τ is a constant having the dimension of time. Calculate the probability that the atom will be found in the 2p state after a sufficiently long time (i.e., $t \to \infty$).

Solution

Since the potential resulting from the interaction of the hydrogen's electron with the external field $\vec{E}(t)$ is $V(t) = -e\vec{r}\cdot\vec{E}(t)$, we can use (10.43) to write the transition probability from the 1s state to 2p as

$$P_{1s\to 2p} = \frac{1}{\hbar^2}\left|\int_{-\infty}^{+\infty}\langle 210 \mid V(t) \mid 100\rangle e^{i\omega_{fi}t}dt\right|^2, \tag{10.210}$$

where

$$\langle 210 \mid V(t) \mid 100\rangle = \langle 210 \mid (-e\vec{r}\cdot\vec{E}) \mid 100\rangle = -\frac{eE_0\tau}{\tau^2+t^2}\langle 210 \mid z \mid 100\rangle. \tag{10.211}$$

Since $z = r\cos\theta$ and

$$\psi_{1s} = R_{10}(r)Y_{00}(\Omega) = \frac{1}{\sqrt{\pi a_0^3}}e^{-r/a_0}, \quad \psi_{2p} = R_{21}(r)Y_{10}(\Omega) = \frac{1}{\sqrt{8\pi a_0^3}}\frac{r}{2a_0}e^{-r/2a_0}\cos\theta, \tag{10.212}$$

and using $\int_0^\pi \sin\theta\cos^2\theta\, d\theta = \int_{-1}^1 x^2 dx = \frac{2}{3}$, we have

$$\begin{aligned}\langle 210 \mid z \mid 100\rangle &= \int_0^\infty r^3 R_{21}^*(r)R_{10}(r)dr \int_0^\pi \sin\theta\cos^2\theta\, d\theta \int_0^{2\pi} d\phi \\ &= \frac{4\pi}{3}\frac{1}{4\pi a_0^4\sqrt{2}}\int_0^\infty r^4 e^{-3r/2a_0}dr = \frac{2^8 a_0}{3^5\sqrt{2}}.\end{aligned} \tag{10.213}$$

Inserting (10.211) and (10.213) into (10.210) we have

$$P_{1s\to 2p} = \frac{2^{15}e^2E_0^2\tau^2a_0^2}{3^{10}\hbar^2}\left|\int_{-\infty}^{+\infty}\frac{e^{i\omega_{fi}t}}{\tau^2+t^2}dt\right|^2. \tag{10.214}$$

We may calculate this integral using the method of residues by closing the contour in the upper half of the t-plane. Since the infinite semicircle has no contribution to the integral, the pole at $t = i\tau$ gives

$$\begin{aligned}\int_{-\infty}^{+\infty}\frac{e^{i\omega_{fi}t}}{\tau^2+t^2}dt &= 2\pi i\,\mathrm{Res}\left[\frac{e^{i\omega_{fi}t}}{\tau^2+t^2}\right]_{t=i\tau} = 2\pi i\lim_{t\to i\tau}\left[\frac{e^{i\omega_{fi}t}}{\tau^2+t^2}\times(t-i\tau)\right] \\ &= 2\pi i\lim_{t\to i\tau}\left[\frac{e^{i\omega_{fi}t}(t-i\tau)}{(t+i\tau)(t-i\tau)}\right] = \frac{\pi}{\tau}e^{-\omega_{fi}\tau},\end{aligned} \tag{10.215}$$

where

$$\omega_{fi} = \frac{1}{\hbar}(E_f - E_i) = \frac{1}{\hbar}\left(E_{2p} - E_{1s}\right) = \frac{1}{\hbar}\left(\frac{1}{4}E_{1s} - E_{1s}\right) = -\frac{3}{4\hbar}E_{1s} = \frac{3R_y}{4\hbar}, \tag{10.216}$$

where R_y is the Rydberg constant: $R_y = 13.6\,\mathrm{eV}$. Inserting (10.215) into (10.214), we obtain the transition probability

$$P_{1s\to 2p} = \frac{2^{15}e^2\pi^2E_0^2a_0^2}{3^{10}\hbar^2}\exp\left(-2\omega_{fi}\tau\right) = \frac{2^{15}e^2\pi^2E_0^2a_0^2}{3^{10}\hbar^2}\exp\left(-\frac{3R_y}{2\hbar}\tau\right). \tag{10.217}$$

Problem 10.9
A hydrogen atom is in its excited 2p state. Calculate the transition rate associated with the 2p → 1s transitions (Lyman) and the lifetime of the 2p state.

Solution

The first expression of the total transition rate is given by (10.144):

$$W_{2p\to 1s} = \frac{4\omega_{2p\to 1s}^3}{3\hbar c^3}\left|\vec{d}_{2p\to 1s}\right|^2, \tag{10.218}$$

where

$$|\vec{d}_{2p\to 1s}|^2 = e^2|\langle 2p \mid \vec{\varepsilon}\cdot\vec{r}\mid 1s\rangle|^2 = e^2\left|\int_0^\infty r^3 R_{21}^* R_{10}(r)\,dr\int d\Omega\, Y_{1m}^*\vec{\varepsilon}\cdot\hat{r}Y_{00}\right|^2. \tag{10.219}$$

First, we need to calculate $\langle 2p \mid \vec{\varepsilon}\cdot\vec{r}\mid 1s\rangle$. The radial integral is given by

$$\int_0^\infty r^3 R_{21}^*(r)R_{10}(r)\,dr = \frac{1}{a_0^4\sqrt{6}}\int_0^\infty r^4 e^{-3r/2a_0}dr = \frac{2^8 a_0}{3^4\sqrt{6}}. \tag{10.220}$$

The angular part can be calculated from (10.130) as follows:

$$\begin{aligned}\int d\Omega Y_{1m}^*(\Omega)\vec{\varepsilon}.\hat{r}Y_{00}(\Omega) &= \sqrt{\frac{4\pi}{3}}\int Y_{1m}^*\left(\frac{-\epsilon_x+i\epsilon_y}{\sqrt{2}}Y_{11}+\frac{\epsilon_x+i\epsilon_y}{\sqrt{2}}Y_{1-1}+\epsilon_z Y_{10}\right)Y_{00}d\Omega\\ &= \frac{1}{\sqrt{3}}\int Y_{1m}^*\left(\frac{-\epsilon_x+i\epsilon_y}{\sqrt{2}}Y_{11}+\frac{\epsilon_x+i\epsilon_y}{\sqrt{2}}Y_{1-1}+\epsilon_z Y_{10}\right)d\Omega\\ &= \frac{1}{\sqrt{3}}\left(\frac{-\epsilon_x+i\epsilon_y}{\sqrt{2}}\delta_{m,1}+\frac{\epsilon_x+i\epsilon_y}{\sqrt{2}}\delta_{m,-1}+\epsilon_z\delta_{m,0}\right),\end{aligned} \tag{10.221}$$

since $\int Y_{1m}^*(\theta,\phi)Y_{l_i m_i}(\theta,\phi)\,d\Omega = \delta_{l_i,1}\delta_{m,m_i}$. An insertion of (10.220) and (10.221) into (10.219) leads to

$$|\vec{d}_{2p\to 1s}|^2 = 32\left(\frac{2}{3}\right)^{10}e^2a_0^2\left[\frac{1}{2}(\epsilon_x^2+\epsilon_y^2)(\delta_{m,-1}+\delta_{m,1})+\epsilon_z^2\delta_{m,0}\right], \tag{10.222}$$

which, when inserted into (10.218), leads to the total transition rate corresponding to a certain value of the azimuthal quantum number m:

$$W_{2p\to 1s} = \frac{4\omega_{2p\to 1s}^3}{3\hbar c^3}|\vec{d}_{fi}|^2 = \frac{128e^2a_0^2\omega^3}{3\hbar c^3}\left(\frac{2}{3}\right)^{10}\left[\frac{1}{2}(\epsilon_x^2+\epsilon_y^2)(\delta_{m,-1}+\delta_{m,1})+\epsilon_z^2\delta_{m,0}\right]. \tag{10.223}$$

Summing over the three possible m-states, $m=-1,0,1$,

$$\sum_{m=-1}^{1}\left[\frac{1}{2}(\epsilon_x^2+\epsilon_y^2)(\delta_{m,-1}+\delta_{m,1})+\epsilon_z^2\delta_{m,0}\right] = \epsilon_x^2+\epsilon_y^2+\epsilon_z^2 = 1, \tag{10.224}$$

and since, as shown in (10.216), $\omega_{2p\to 1s} = (E_{2p}-E_{1s})/\hbar = 3R_y/(4\hbar) = 3e^2/(8\hbar a_0)$ (because the Rydberg constant R_y is equal to $e^2/(2a_0)$), we can reduce (10.223) to

$$W_{2p\to 1s} = \frac{128}{3\hbar c^3}\left(\frac{2}{3}\right)^{10}e^2a_0^2\omega_{2p\to 1s}^3 = \left(\frac{2}{3}\right)^8\left(\frac{e^2}{\hbar c}\right)^4\frac{c}{a_0} = \left(\frac{2}{3}\right)^8\frac{c\alpha^4}{a_0}, \tag{10.225}$$

where $\alpha = e^2/(\hbar c) = 1/137$ is the fine structure constant and $a_0 = 0.529 \times 10^{-10}\,\mathrm{m}$ is the Bohr radius. The numerical value of the transition rate is

$$W_{2p\to 1s} = \left(\frac{2}{3}\right)^8 \frac{c\alpha^4}{a_0} \simeq \left(\frac{2}{3}\right)^8 \frac{3\times 10^8\,\mathrm{m\,s^{-1}}}{137^4 \times 0.529\times 10^{-10}\,\mathrm{m}} = 0.628\times 10^9\,\mathrm{s^{-1}}. \tag{10.226}$$

The lifetime of the 2p state is then given by

$$\tau = \frac{1}{W_{2p\to 1s}} = \left(\frac{3}{2}\right)^8 \frac{a_0}{c\alpha^4} = \frac{1.5^8 \times 137^4 \times 0.529\times 10^{-10}\,\mathrm{m}}{3\times 10^8\,\mathrm{m\,s^{-1}}} = 1.6\times 10^{-9}\,\mathrm{s}. \tag{10.227}$$

This value is in very good agreement with experimental data.

Remark

Another way of obtaining (10.225) is to use the relation

$$\begin{aligned} W_{2p\to 1s} &= \frac{4e^2\omega_{2p\to 1s}^3}{3\hbar c^3}\frac{1}{3}\sum_{m=-1}^{1} |\langle 21m \mid \vec{r} \mid 100\rangle|^2 \\ &= \frac{4e^2\omega_{2p\to 1s}^3}{9\hbar c^3}\sum_{m=-1}^{1}\Big[|\langle 21m \mid \hat{x} \mid 100\rangle|^2 + |\langle 21m \mid \hat{y} \mid 100\rangle|^2 + |\langle 21m \mid \hat{z} \mid 100\rangle|^2\Big], \end{aligned} \tag{10.228}$$

where we have averaged over the various transitions. Using the relations $x = r\sin\theta\cos\phi = -\sqrt{2\pi/3}\,r(Y_{11} - Y_{1-1})$, $y = r\sin\theta\sin\phi = i\sqrt{2\pi/3}\,r(Y_{11} + Y_{1-1})$, and $z = r\cos\theta = \sqrt{4\pi/3}\,rY_{10}$, we can show that

$$\begin{aligned} \langle 21m \mid \hat{x} \mid 100\rangle &= -\frac{1}{\sqrt{4\pi}}\sqrt{\frac{2\pi}{3}}\int_0^\infty r^3 R_{21}^*(r)R_{10}(r)dr \int Y_{1m}^*(\Omega)\,(Y_{11} - Y_{1-1})\,d\Omega \\ &= -\frac{1}{\sqrt{6}}\left[\frac{24}{\sqrt{6}}\left(\frac{2}{3}\right)^5 a_0\right](\delta_{m,1} - \delta_{m,-1}), \end{aligned} \tag{10.229}$$

$$\begin{aligned} \langle 21m \mid \hat{y} \mid 100\rangle &= \frac{i}{\sqrt{4\pi}}\sqrt{\frac{2\pi}{3}}\int_0^\infty r^3 R_{21}^*(r)R_{10}(r)dr \int Y_{1m}^*(Y_{11} + Y_{1-1})d\Omega \\ &= \frac{i}{\sqrt{6}}\left[\frac{24}{\sqrt{6}}\left(\frac{2}{3}\right)^5 a_0\right](\delta_{m,1} + \delta_{m,-1}), \end{aligned} \tag{10.230}$$

$$\begin{aligned} \langle 21m \mid \hat{z} \mid 100\rangle &= \frac{1}{\sqrt{4\pi}}\sqrt{\frac{4\pi}{3}}\int_0^\infty r^3 R_{21}^*(r)R_{10}(r)dr \int Y_{1m}^* Y_{10}\,d\Omega \\ &= \frac{1}{\sqrt{3}}\left[\frac{24}{\sqrt{6}}\left(\frac{2}{3}\right)^5 a_0\right]\delta_{m,0}. \end{aligned} \tag{10.231}$$

A combination of the previous three relations leads to

$$\begin{aligned}\sum_{m=-1}^{1}|\langle 21m \mid \vec{r} \mid 100\rangle|^2 &= 96a_0^2\left(\frac{2}{3}\right)^{10}\sum_m\left[\frac{1}{6}(\delta_{m,1}-\delta_{m,-1})^2+\frac{1}{6}(\delta_{m,1}+\delta_{m,-1})^2+\frac{1}{3}\delta_{m,0}^2\right]\\ &= 96a_0^2\left(\frac{2}{3}\right)^{10}\sum_m\left[\frac{1}{6}(\delta_{m,1}+\delta_{m,-1})+\frac{1}{6}(\delta_{m,1}+\delta_{m,-1})+\frac{1}{3}\delta_{m,0}\right]\\ &= \frac{96a_0^2}{3}\left(\frac{2}{3}\right)^{10}\sum_{m=-1}^{1}(\delta_{m,-1}+\delta_{m,1}+\delta_{m,0}) = 96\left(\frac{2}{3}\right)^{10}a_0^2.\end{aligned} \tag{10.232}$$

Finally, substituting (10.232) into (10.228) and using $\omega_{2p\to 1s} = 3e^2/(8\hbar a_0)$, we obtain

$$W_{2p\to 1s} = \frac{128e^2a_0^2}{3\hbar c^3}\omega^3\left(\frac{2}{3}\right)^{10} = \left(\frac{2}{3}\right)^8\left(\frac{e^2}{\hbar c}\right)^4\frac{c}{a_0} = \left(\frac{2}{3}\right)^8\frac{c\alpha^4}{a_0}. \tag{10.233}$$

Problem 10.10

(a) Calculate the transition rate from the first excited state to the ground state for an isotropic (three-dimensional) harmonic oscillator of charge q.

(b) Find a numerical value for the rate calculated in (a) as well as the lifetime of the first excited state for the case of an electron (i.e., $m_ec^2 = 0.511\,\text{MeV}$) oscillating with a frequency of an optical radiation $\omega \simeq 10^{15}\,\text{rad}\,\text{s}^{-1}$.

Solution

As mentioned in Chapter 6, the ground state of an isotropic harmonic oscillator is a 1s state, $(n,l,m) = (0,0,0)$, whose energy and wave function are $E_0 = 3\hbar\omega/2$ and

$$\psi_{000}(r,\,\theta,\,\phi) = R_{00}(r)Y_{00}(\theta,\,\phi) = \frac{2}{\sqrt{\sqrt{\pi}}}\left(\frac{m\omega}{\hbar}\right)^{3/4}e^{-m\omega r^2/2\hbar}Y_{00}(\theta,\,\phi), \tag{10.234}$$

and the first excited state is a $1p$ state $(n,l,m) = (1,1,m)$ whose energy and wave function are $E_1 = 5\hbar\omega/2$ and

$$\psi_{11m}(r,\,\theta,\,\phi) = R_{11}(r)Y_{1m}(\theta,\,\phi) = \sqrt{\frac{8}{3\sqrt{\pi}}}\left(\frac{m\omega}{\hbar}\right)^{5/4}re^{-m\omega r^2/2\hbar}Y_{1m}(\theta,\,\phi). \tag{10.235}$$

Using $\int_0^\infty x^4e^{-x^2}dx = \frac{3}{8}\sqrt{\pi}$ along with a change of variable $x = \sqrt{m\omega/\hbar}\,r$, we have

$$\int_0^\infty r^3R_{11}^*(r)R_{10}(r)\,dr = 4\sqrt{\frac{2}{3\pi}}\left(\frac{m\omega}{\hbar}\right)^2\int_0^\infty r^4e^{-m\omega r^2/\hbar}\,dr = \sqrt{\frac{3\hbar}{2m\omega}}. \tag{10.236}$$

(a) The transition rate for a $1p \to 1s$ transition is given by

$$W_{1p\to 1s} = \frac{4q^2\omega_{1p\to 1s}^3}{3\hbar c^3}\frac{1}{3}\sum_{m=-1}^{1}|\langle 11m \mid \vec{r} \mid 000\rangle|^2$$
$$= \frac{4q^2\omega_{1p\to 1s}^3}{9\hbar c^3}\sum_{m=-1}^{1}\Big[|\langle 11m \mid \hat{x} \mid 000\rangle|^2 + |\langle 11m \mid \hat{y} \mid 000\rangle|^2 + |\langle 21m \mid \hat{z} \mid 000\rangle|^2\Big]. \tag{10.237}$$

Since $x = r\sin\theta\cos\phi = -\sqrt{2\pi/3}\,r(Y_{11} - Y_{1-1})$, $y = r\sin\theta\sin\phi = i\sqrt{2\pi/3}\,r(Y_{11} + Y_{1-1})$, and $z = r\cos\theta = \sqrt{4\pi/3}\,rY_{10}$, and using (10.236), we can show by analogy with (10.229) to (10.231) that

$$\langle 11m \mid \hat{x} \mid 000\rangle = -\frac{1}{\sqrt{4\pi}}\sqrt{\frac{2\pi}{3}}\int_0^\infty r^3R_{11}^*(r)R_{00}(r)\,dr\int Y_{1m}^*(\Omega)\,(Y_{11} - Y_{1-1})\,d\Omega$$
$$= -\frac{1}{\sqrt{6}}\sqrt{\frac{3\hbar}{2m\omega}}(\delta_{m,1} - \delta_{m,-1}), \tag{10.238}$$

$$\langle 11m \mid \hat{y} \mid 000\rangle = \frac{i}{\sqrt{4\pi}}\sqrt{\frac{2\pi}{3}}\int_0^\infty r^3R_{11}^*(r)R_{00}(r)\,dr\int Y_{1m}^*(Y_{11} + Y_{1-1})\,d\Omega$$
$$= \frac{i}{\sqrt{6}}\sqrt{\frac{3\hbar}{2m\omega}}(\delta_{m,1} + \delta_{m,-1}), \tag{10.239}$$

$$\langle 11m \mid \hat{z} \mid 000\rangle = \frac{1}{\sqrt{4\pi}}\sqrt{\frac{4\pi}{3}}\int_0^\infty r^3R_{11}^*(r)R_{00}(r)\,dr\int Y_{1m}^*Y_{10}\,d\Omega = \frac{1}{\sqrt{3}}\sqrt{\frac{3\hbar}{2m\omega}}\delta_{m,0}. \tag{10.240}$$

A combination of the previous three relations leads to

$$\sum_{m=-1}^{1}|\langle 11m \mid \vec{r} \mid 000\rangle|^2 = \frac{3\hbar}{2m\omega}\sum_m\left[\frac{1}{6}(\delta_{m,1} - \delta_{m,-1})^2 + \frac{1}{6}(\delta_{m,1} + \delta_{m,-1})^2 + \frac{1}{3}\delta_{m,0}^2\right]$$
$$= \frac{\hbar}{2m\omega}\sum_{m=-1}^{1}(\delta_{m,-1} + \delta_{m,1} + \delta_{m,0}) = \frac{3\hbar}{2m\omega}. \tag{10.241}$$

Substituting (10.241) into (10.237), and using $\omega_{1p\to 1s} = (E_1 - E_0)/\hbar = (\frac{5}{2} - \frac{3}{2})\omega = \omega$, we obtain

$$W_{1p\to 1s} = \frac{4q^2\omega_{1p\to 1s}^3}{9\hbar c^3}\frac{3\hbar}{2m\omega} = \frac{2q^2\omega^2}{3mc^3}. \tag{10.242}$$

(b) In the case of an electron ($q = -e$ and $m_ec^2 = 0.511\,\text{MeV}$) which is oscillating with a frequency of $\omega \simeq 10^{15}\,\text{s}^{-1}$, the transition rate is

$$W_{1p\to 1s} = \frac{2e^2\omega^2}{3m_ec^3} = \frac{2\alpha}{3}\left(\frac{\hbar c}{m_ec^2}\right)\frac{\omega^2}{c}$$
$$= \frac{2}{3}\frac{1}{137}\left(\frac{197\,\text{MeV fm}}{0.511\,\text{MeV}}\right)\frac{10^{30}\,\text{s}^{-2}}{3\times 10^8\,\text{m s}^{-1}} \simeq 0.64\times 10^7\,\text{s}^{-1}, \tag{10.243}$$

where $\alpha = e^2/(\hbar c) = 1/137$ is the fine structure constant. The lifetime of the 1p state for the oscillator is given by

$$\tau = \frac{1}{W_{1p\to 1s}} = \frac{3m_e c^3}{2e^2\omega^2} \simeq \frac{1}{0.64\times 10^7\,s^{-1}} = 1.56\times 10^{-7}\,s. \tag{10.244}$$

Problem 10.11
Show that free electrons can neither emit nor absorb photons.

Solution
If the electron is *free* both before and after it interacts with the photon, its initial and final wave functions are given by *plane* waves: $\psi_i(\vec{r}) = (2\pi)^{-3/2}e^{i\vec{k}_i\cdot\vec{r}}$ and $\psi_f(\vec{r}) = (2\pi)^{-3/2}e^{i\vec{k}_f\cdot\vec{r}}$. Let us assume, for argument sake, that a free electron can absorb and emit a photon; the corresponding absorption and emission transition rates would be given as follows (see (10.97) and (10.98)):

$$\Gamma^{abs}_{i\to f} = \frac{4\pi^2e^2}{m_e^2\omega V}\left|(\vec{\varepsilon}\cdot\vec{k}_i)\langle\psi_f\mid e^{i\vec{k}\cdot\vec{r}}\mid\psi_i\rangle\right|^2\delta(E_f - E_i - \hbar\omega), \tag{10.245}$$

$$\Gamma^{emi}_{i\to f} = \frac{4\pi^2e^2}{m_e^2\omega V}\left|(\vec{\varepsilon}\cdot\vec{k}_i)\langle\psi_f\mid e^{-i\vec{k}\cdot\vec{r}}\mid\psi_i\rangle\right|^2\delta(E_f - E_i + \hbar\omega), \tag{10.246}$$

where we have used $\vec{P}\psi_i(\vec{r}) = \vec{k}_i\psi_i(\vec{r})$. Since

$$\langle\psi_f\mid e^{\pm i\vec{k}\cdot\vec{r}}\mid\psi_i\rangle = \frac{1}{(2\pi)^2}\int d^3r\, e^{i(\vec{k}_i-\vec{k}_f\pm\vec{k})\cdot\vec{r}} = \delta(\vec{k}_i - \vec{k}_f \pm \vec{k}), \tag{10.247}$$

the delta functions $\delta(\vec{k}_i - \vec{k}_f \pm \vec{k})$ give the conservation laws of the linear momentum for both the absorption and emission processes.

Let us show first that a free electron cannot absorb a photon. For this, we are going to show that the momentum conservation condition $\delta(\vec{k}_i - \vec{k}_f + \vec{k})$ is incompatible with the energy conservation condition $\delta(E_f - E_i - \hbar\omega)$. Combining equations (10.245) and (10.247), we see that the absorption rate is proportional to the product of two delta functions: $\Gamma^{abs}_{i\to f} \sim \delta(\vec{k}_i - \vec{k}_f + \vec{k})\delta(E_f - E_i - \hbar\omega)$, one pertaining to the conservation of momentum

$$\delta(\vec{k}_i - \vec{k}_f + \vec{k}) \quad\Longrightarrow\quad \vec{p}_i - \vec{p}_f + \vec{p}_{photon} = 0, \tag{10.248}$$

the other dealing with the conservation of energy

$$\delta(E_f - E_i - \hbar\omega) \quad\Longrightarrow\quad E_f - E_i - cp_{photon} = 0, \tag{10.249}$$

where $\vec{p}_i = \hbar\vec{k}_i$ and E_i are the initial momentum and energy of the electron, $\vec{p}_f = \hbar\vec{k}_f$ and E_f are its final momentum and energy, and $\vec{p}_{photon} = \hbar\vec{k}$ and cp_{photon} are the linear momentum and energy of the absorbed photon. We are now ready to show that the condition (10.248) is incompatible with (10.249). If we work within the rest frame of the initial electron, we have $\vec{p}_i = 0$. Thus, on the one hand, (10.248) leads to $\vec{p}_{photon} = \vec{p}_f$ and, on the other hand, (10.249) leads to $E_f = cp_{photon}$ or $p_f^2/2m_e = cp_{photon}$. Indeed, conditions (10.248) and (10.249) are

contradictory since, inserting $\vec{p}_i = 0$ and $\vec{p}_{photon} = \vec{p}_f$ into (10.249), we end up with $p_f^2/2m_e = cp_f$ or $p_f = 2m_e c$. This suggests either that $v_f = 0$ and this is meaningless since, as $\vec{p}_{photon} = \vec{p}_f$, the speed of the photon would also be zero; or that $v_f = 2c$ and this is impossible. So both results are impossible. In summary, having started with the assumption that a free electron can absorb a photon (10.245), we have ended up with a momentum conservation law and an energy conservation law that are contradictory. Thus, a free electron cannot absorb a photon.

Following the same procedure, we can also show that the assumption of a free electron emitting a photon leads to a momentum conservation law and an energy conservation law that are incompatible; thus, a free electron cannot emit a photon.

Problem 10.12

A hydrogen atom in its ground state is placed in an oscillating electric field $\vec{\mathcal{E}}(t) = \vec{\mathcal{E}}_0 \sin(\omega t)$ of angular frequency ω with $\hbar\omega > m_e e^4/(2\hbar^2)$.

(a) Find the transition rate (probability per unit time) that the atom will be ionized.

(b) Use the expression derived in (a) to find the maximum transition rate.

Solution

After ionization we assume the electron to be in free motion: its energy is purely kinetic $E_k = \hbar^2 k^2/2m_e$ and its wave function is a plane wave $\psi_k(\vec{r}) = (2\pi)^{-3/2} e^{i\vec{k}\cdot\vec{r}}$. Since the perturbation resulting from the interaction of the hydrogen's electron with the external field $\vec{\mathcal{E}}(t)$ is *harmonic*,

$$\hat{V}(t) = -e\vec{r}\cdot\vec{\mathcal{E}}(t) = -e\vec{r}\cdot\vec{\mathcal{E}}_0 \sin(\omega t) = \frac{e}{2i}\vec{r}\cdot\vec{\mathcal{E}}_0 e^{-i\omega t} - \frac{e}{2i}\vec{r}\cdot\vec{\mathcal{E}}_0 e^{i\omega t}, \tag{10.250}$$

we can infer, by analogy with the method that led to (10.56) from (10.52), the transition rate for the ionization of the hydrogen atom:

$$\begin{aligned}\Gamma_{0k} &= \frac{2\pi}{\hbar}\left|\frac{e}{2i}\langle\psi_k|\vec{r}\cdot\vec{\mathcal{E}}_0 \mid 100\rangle\right|^2 \delta(E_k - E_0 + \hbar\omega) \\ &\quad + \frac{2\pi}{\hbar}\left|-\frac{e}{2i}\langle\psi_k|\vec{r}\cdot\vec{\mathcal{E}}_0 \mid 100\rangle\right|^2 \delta(E_k - E_0 - \hbar\omega),\end{aligned} \tag{10.251}$$

where $E_0 = -m_e e^4/2\hbar^2 = -13.6\,\text{eV}$ is the ground state energy and $E_k = \hbar^2 k^2/2m_e$ is the final energy of the electron. The first delta term, $\delta(E_k - E_0 + \hbar\omega)$, in (10.251) does not contribute, since if $\hbar\omega = E_0 - E_k$ the ionization could not take place because the electric field would not be strong enough to ionize the atom. The transition rate (10.251) then becomes

$$\Gamma_{0k} = \frac{\pi e^2}{2\hbar}\left|\langle\psi_k|\vec{r}\cdot\vec{\mathcal{E}}_0 \mid 100\rangle\right|^2 \delta(E_k - E_0 - \hbar\omega). \tag{10.252}$$

To calculate $\langle\psi_k|\vec{r}\cdot\vec{\mathcal{E}}_0 \mid 100\rangle$, let us take $\vec{k}$ along the z-axis and hence $\vec{k}\cdot\vec{r} = kr\cos\theta$ and $\psi_k(\vec{r}) = (2\pi)^{-3/2} e^{ikr\cos\theta}$. Taking (θ, ϕ) and (α, β) as the respective polar angles of $\vec{r}$ and $\vec{\mathcal{E}}_0$, we have $\vec{r} = r(\sin\theta\cos\phi\,\hat{\imath} + \sin\theta\sin\phi\,\hat{\jmath} + \cos\theta\,\hat{k})$ and $\vec{\mathcal{E}}_0 = \mathcal{E}_0(\sin\alpha\cos\beta\,\hat{\imath} + \sin\alpha\sin\beta\,\hat{\jmath} + \cos\alpha\,\hat{k})$; hence

$$\begin{aligned}\vec{r}\cdot\vec{\mathcal{E}}_0 &= r\mathcal{E}_0(\sin\theta\cos\phi\sin\alpha\cos\beta + \sin\theta\sin\phi\sin\alpha\sin\beta + \cos\theta\cos\alpha) \\ &= r\mathcal{E}_0\Big[\sin\theta\sin\alpha\cos(\phi - \beta) + \cos\theta\cos\alpha\Big].\end{aligned} \tag{10.253}$$

Since $\psi_{1s} = (\pi a_0^3)^{-1/2} e^{-r/a_0}$ and $d^3r = r^2 dr \sin\theta \cos\phi \, d\theta \, d\phi$, we have

$$\begin{aligned}\langle \psi_k | \vec{r} \cdot \vec{\mathcal{E}}_0 \mid 100 \rangle &= \frac{1}{(2\pi)^{3/2}} \frac{1}{\sqrt{\pi a_0^3}} \int d^3r \, (\vec{r} \cdot \vec{\mathcal{E}}_0) e^{-ikr\cos\theta - r/a_0} \\ &= \frac{\mathcal{E}_0}{\sqrt{8\pi^4 a_0^3}} \int_0^\infty r^3 e^{-r/a_0} dr \int_0^\pi \sin\theta e^{-ikr\cos\theta} d\theta \int_0^{2\pi} \Big[\sin\theta \sin\alpha \cos(\phi - \beta) + \cos\theta \cos\alpha \Big] d\phi \\ &= \frac{2\pi \mathcal{E}_0 \cos\alpha}{\sqrt{8\pi^4 a_0^3}} \int_0^\infty r^3 e^{-r/a_0} \, dr \int_0^\pi \sin\theta \cos\theta e^{-ikr\cos\theta} \, d\theta, \end{aligned} \tag{10.254}$$

where we have used $\int_0^{2\pi} \cos(\phi - \beta) d\phi = 0$, since $\int_0^{2\pi} \cos\phi \, d\phi = 0$ and $\int_0^{2\pi} \sin\phi \, d\phi = 0$. A change of variable $x = \cos\theta$ and an integration by parts leads to

$$\begin{aligned}\int_0^\pi \sin\theta \cos\theta e^{-ikr\cos\theta} d\theta &= \int_{-1}^1 x e^{-ikrx} dx = \frac{1}{-ikr} x e^{-ikrx} \Big|_{-1}^1 - \frac{1}{(-ikr)^2} e^{-ikrx} \Big|_{-1}^1 \\ &= \frac{i}{kr} \Big[e^{-ikr} + e^{ikr} \Big] + \frac{1}{k^2 r^2} \Big[e^{-ikr} - e^{ikr} \Big]. \end{aligned} \tag{10.255}$$

When we insert this integral into (10.254), we still need to calculate four radial integrals which can be carried out by parts:

$$\int_0^\infty r e^{\pm ikr - r/a_0} dr = \frac{1}{\pm ik - 1/a_0} r e^{\pm ikr - r/a_0} \Big|_0^\infty - \frac{1}{(\pm ik - 1/a_0)^2} e^{\pm ikr - r/a_0} \Big|_0^\infty = \frac{a_0^2}{(\pm i a_0 k - 1)^2}, \tag{10.256}$$

$$\begin{aligned}\int_0^\infty r^2 e^{\pm ikr - r/a_0} dr &= \frac{1}{\pm ik - 1/a_0} r^2 e^{\pm ikr - r/a_0} \Big|_0^\infty - \frac{2}{\pm ik - 1/a_0} \int_0^\infty r e^{\pm ikr - r/a_0} dr \\ &= -\frac{2}{(\pm ik - 1/a_0)^2} r e^{\pm ikr - r/a_0} \Big|_0^\infty + \frac{2}{(\pm ik - 1/a_0)^3} e^{\pm ikr - r/a_0} \Big|_0^\infty \\ &= -\frac{2a_0^3}{(\pm i a_0 k - 1)^3}. \end{aligned} \tag{10.257}$$

Inserting (10.255) to (10.257) into (10.254), we obtain

$$\begin{aligned}\langle \psi_k | \vec{r} \cdot \vec{\mathcal{E}}_0 \mid 100 \rangle &= \frac{2\pi \mathcal{E}_0 \cos\alpha}{\sqrt{8\pi^4 a_0^3}} \left[\frac{a_0^2}{k^2(-ia_0k - 1)^2} - \frac{a_0^2}{k^2(ia_0k - 1)^2} - \frac{2ia_0^3}{k(ia_0k - 1)^3} \right. \\ &\qquad \left. - \frac{2ia_0^3}{k(-ia_0k - 1)^3} \right] \\ &= -\frac{16\mathcal{E}_0 \cos\alpha}{\pi\sqrt{2a_0^5}} \frac{ia_0^6 k}{(a_0^2k^2 + 1)^3}. \end{aligned} \tag{10.258}$$

A substitution of this expression into (10.252) leads to

$$\Gamma_{0k} = \frac{\pi e^2}{2\hbar} \frac{128 \mathcal{E}_0^2 \cos^2\alpha}{\pi^2 a_0^5} \frac{k^2 a_0^{12}}{(a_0^2k^2 + 1)^6} \delta(E_k - E_0 + \hbar\omega). \tag{10.259}$$

This relation gives the transition rate for a single final state ψ_k corresponding to a given k. We need to sum over all final states of the electron. These represent a continuum; we must then integrate over all directions of emission and over all possible momenta:

$$\begin{aligned}\Gamma_0 &= \int \Gamma_{0k}\, d^3k = \int k^2 dk \int_0^\pi \Gamma_{0k} \sin\alpha\, d\alpha \int_0^{2\pi} d\beta \\ &= 2\pi \frac{64e^2\mathcal{E}_0^2 a_0^7}{\pi\hbar} \int \frac{k^4 \delta(E_k - E_0 - \hbar\omega)}{(a_0^2k^2+1)^6} dk \int_0^\pi \sin\alpha\cos^2\alpha\, d\alpha \\ &= \frac{256e^2\mathcal{E}_0^2 a_0^7}{3\hbar} \int \frac{k^4 \delta(E_k - E_0 - \hbar\omega)}{(a_0^2k^2+1)^6} dk, \end{aligned} \tag{10.260}$$

where we have used $\int_0^\pi \sin\alpha\cos^2\alpha\, d\alpha = \int_{-1}^1 x^3 dx = \frac{2}{3}$. The integration over k can be converted into an integration over the final energy E_k: since $E_k = \hbar^2k^2/(2m_e)$, a change of variable $k = \sqrt{2m_eE_k/\hbar^2}$, and hence $k\,dk = (m_e/\hbar^2)dE_k$, reduces (10.260) to

$$\begin{aligned}\Gamma_0 &= \frac{256e^2\mathcal{E}_0^2 a_0^7}{3\hbar} \int \frac{k^3 \delta(E_k - E_0 - \hbar\omega)}{(a_0^2k^2+1)^6} k\, dk \\ &= \frac{m_e}{\hbar^2}\frac{256e^2\mathcal{E}_0^2 a_0^7}{3\hbar} \int \frac{(2m_eE_k/\hbar^2)^{3/2}\delta(E_k - E_0 - \hbar\omega)}{\left(2m_ea_0^2E_k/\hbar^2+1\right)^6} dE_k \\ &= \frac{256e^2m_e\mathcal{E}_0^2 a_0^7}{3\hbar^3} \frac{(2m_e/\hbar^2)^{3/2}(E_0+\hbar\omega)^{3/2}}{\left[2m_ea_0^2(E_0+\hbar\omega)/\hbar^2+1\right]^6}. \end{aligned} \tag{10.261}$$

This relation can be simplified if we use $E_0 = -m_ee^4/(2\hbar^2) = -\hbar\omega_0$, which gives $E_0 + \hbar\omega = \hbar(\omega - \omega_0) = \hbar\omega_0(\omega/\omega_0 - 1)$. Since $a_0 = \hbar^2/(m_ee^2)$, we have $\hbar\omega_0 a_0^2 = m_ee^4\hbar^4/(2\hbar^2m_e^2e^4) = \hbar^2/(2m_e)$ and hence $2m_ea_0^2(E_0+\hbar\omega)/\hbar^2 = 2m_e\hbar\omega_0a_0^2(\omega/\omega_0 - 1)/\hbar^2 = \omega/\omega_0 - 1$. Thus, inserting the expressions $E_0 + \hbar\omega = \hbar\omega_0(\omega/\omega_0 - 1)$ and $2m_ea_0^2(E_0+\hbar\omega)/\hbar^2 + 1 = \omega/\omega_0$ into (10.261), we obtain

$$\Gamma_0 = \frac{256e^2m_e\mathcal{E}_0^2 a_0^7}{3\hbar^3} \frac{(2m_e/\hbar^2)^{3/2}(\hbar\omega_0)^{3/2}(\omega/\omega_0 - 1)^{3/2}}{(\omega/\omega_0)^6}. \tag{10.262}$$

Finally, since $(2m_e/\hbar^2)^{3/2}(\hbar\omega_0)^{3/2} = (2m_e/\hbar^2)^{3/2}(m_ee^4/2\hbar^2)^{3/2} = m_e^3e^6/\hbar^6$ and using $a_0^4 = \hbar^8/(m^4e^8)$, we can write (10.262) as

$$\Gamma_0 = \frac{256\mathcal{E}_0^2 a_0^3}{3\hbar}\left(\frac{\omega_0}{\omega}\right)^6\left(\frac{\omega}{\omega_0} - 1\right)^{3/2}. \tag{10.263}$$

If the frequency of the oscillating electric field is smaller than or equal to ω_0, the atom will not be ionized; at $\omega = \omega_0$ the probability of ionization will be zero.

(b) The maximum transition rate is obtained by taking the derivative of (10.263):

$$\frac{d\Gamma_0}{d\omega} = 0 \Longrightarrow \frac{2}{\omega}\left(\frac{\omega}{\omega_0} - 1\right) = \frac{1}{2\omega_0} \Longrightarrow \omega = \frac{4}{3}\omega_0. \tag{10.264}$$

Inserting $\omega = \frac{4}{3}\omega_0$ into (10.263) we obtain the maximum transition rate

$$\Gamma_{0max} = \frac{256\mathcal{E}_0^2 a_0^3}{3\hbar}\left(\frac{3}{4}\right)^6\left(\frac{4}{3} - 1\right)^{3/2} = \frac{\mathcal{E}_0^2 a_0^3}{\hbar}\frac{3^{7/2}}{2^4}. \tag{10.265}$$

10.7 Exercises

Exercise 10.1
Consider a spinless particle of mass m in a one-dimensional infinite potential well with walls at $x = 0$ and $x = a$ which is initially (i.e., at $t = 0$) in the state $\psi(x,0) = [\phi_1(x) + \phi_3(x)]/\sqrt{2}$, where $\phi_1(x)$ and $\phi_3(x)$ are the ground and second excited states, respectively, with $\phi_n(x) = \sqrt{2/a}\sin(n\pi x/a)$.

(a) What is the state vector $\psi(x,t)$ for $t > 0$ in the Schrödinger picture.

(b) Evaluate $\langle\hat{X}\rangle$, $\langle\hat{P}\rangle$, $\langle\hat{X}^2\rangle$, and $\langle\hat{P}^2\rangle$ as functions of time for $t > 0$ in the Schrödinger picture.

(c) Repeat part (b) in the Heisenberg picture: i.e., evaluate $\langle\hat{X}\rangle_H$, $\langle\hat{P}\rangle_H$, $\langle\hat{X}^2\rangle_H$, and $\langle\hat{P}^2\rangle_H$ as functions of time for $t > 0$.

Exercise 10.2
Evaluate the expectation value $\langle\hat{X}_H(t)\hat{P}\rangle_3$ for the third excited state of a one-dimensional harmonic oscillator.

Exercise 10.3
Evaluate the expectation value $\langle\hat{X}\hat{P}_H(t)\rangle_n$ for the nth excited state of a one-dimensional harmonic oscillator.

Exercise 10.4
Consider a one-dimensional harmonic oscillator which is initially (i.e., at $t = 0$) in the state $|\psi(0)\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, where $|0\rangle$ and $|1\rangle$ are the ground and first excited states, respectively.

(a) What is the state vector $|\psi(t)\rangle$ for $t > 0$ in the Schrödinger picture?

(b) Evaluate $\langle\hat{X}\rangle$, $\langle\hat{P}\rangle$, $\langle\hat{X}^2\rangle$, and $\langle\hat{P}^2\rangle$ as functions of time for $t > 0$ in the Schrödinger picture.

(c) Repeat part (b) in the Heisenberg picture.

Exercise 10.5

(a) Calculate the coordinate operator $\hat{X}_H(t)$ for a free particle in one dimension in the Heisenberg picture.

(b) Evaluate the commutator $[\hat{X}_H(t),\ \hat{X}_H(0)]$.

Exercise 10.6
Consider the Hamiltonian $H = -(eB/mc)\hat{S}_x = \omega\hat{S}_x$.

(a) Write down the Heisenberg equations of motion for the time-dependent operators $\hat{S}_x(t)$, $\hat{S}_y(t)$, and $\hat{S}_z(t)$.

(b) Solve these equations to obtain S_x, S_y, S_z as functions of time.

Exercise 10.7
Evaluate the quantity $\langle n \mid \hat{P}_H(t)\hat{P} \mid n\rangle$ for the nth excited state of a one-dimensional harmonic oscillator, where $\hat{P}_H(t)$ and $\hat{P}$ designate the momentum operators in the Heisenberg picture and the Schrödinger picture, respectively.

Exercise 10.8
The Hamiltonian due to the interaction of a particle of mass m, charge q (the charge is negative), and spin $\vec{S}$ with a magnetic field pointing along the y-axis is $\hat{H} = -(qB/mc)\hat{S}_y$.

(a) Use the Heisenberg equation to calculate $d\hat{S}_x/dt$, $d\hat{S}_y/dt$, and $d\hat{S}_z/dt$.

(b) Solve these equations to obtain the components of the spin operator as functions of time.

Exercise 10.9
A particle is initially (i.e., when $t < 0$) in its ground state in a one-dimensional harmonic oscillator potential. At $t = 0$ a perturbation $\hat{V}(x,t) = V_0\hat{x}^2e^{-t/\tau}$ is turned on. Calculate to first order the probability that, after a sufficiently long time (i.e., $t \to \infty$), the system will have made a transition to a given excited state; consider all final states.

Exercise 10.10
A particle, initially (i.e., when $t < 0$) in its ground state in an infinite potential well whose walls are located at $x = 0$ and $x = a$, is subject, starting at time $t = 0$, to a time-dependent perturbation $\hat{V}(t) = V_0\hat{x}^2e^{-t^2}$ where V_0 is a small parameter. Calculate the probability that the particle will be found in its second excited state at $t = +\infty$.

Exercise 10.11
Find the intensity associated with the transition 3s $\to$ 2p in the hydrogen atom.

Exercise 10.12
A hydrogen atom in its ground state is placed in a region where, at $t = 0$, a time-dependent electric field is turned on:

$$\vec{E}(t) = E_0(\hat{\imath} + \hat{\jmath} + \hat{k})e^{-t/\tau},$$

where τ is a positive real number. Using first-order time-dependent perturbation theory, calculate the probability that, after a sufficiently long time (i.e., $t \gg \tau$), the atom is to be found in each of the $n = 2$ states (i.e., consider the transitions to all the states in the $n = 2$ level). *Hint:* You may use: $\int_0^\infty r^3R_{21}^*(r)R_{10}(r)\,dr = (24a_0/\sqrt{6})\left(\frac{2}{3}\right)^5$.

Exercise 10.13
(a) Calculate the reduced matrix element $\langle 1 \parallel Y_1 \parallel 2\rangle$. *Hint:* For this, you may need to calculate $\langle 1, 0|Y_{10}|2, 0\rangle$ directly and then from the Wigner–Eckart theorem.

(b) Using the Wigner–Eckart theorem and the relevant Clebsch–Gordan coefficients from tables, calculate $\langle 1, m|Y_{1m'}|2, m''\rangle$ for all possible values of m, m', and m''.

(c) Using the results of part (b), calculate the 3d $\to$ 2p transition rate for the hydrogen atom in the dipole approximation; give a numerical value. *Hint:* You may use the integral $\int_0^\infty r^3R_{21}^*(r)R_{32}(r)\,dr = (64a_0/15\sqrt{5})\left(\frac{6}{5}\right)^5$ and the following Clebsch–Gordan coefficients:

$$\langle j, 1;\ m, 0|(j-1),\ m\rangle = -\sqrt{(j-m)(j+m)/[j(2j+1)]},$$
$$\langle j, 1;\ (m-1), 1|(j-1),\ m\rangle = \sqrt{(j-m)(j-m+1)/[2j(2j+1)]}, \text{ and}$$
$$\langle j, 1;\ (m+1), -1|(j-1),\ m\rangle = \sqrt{(j+m)(j+m+1)/[2j(2j+1)]}.$$

Exercise 10.14
A particle is initially in its ground state in an infinite one-dimensional potential box with sides

at $x = 0$ and $x = a$. If the wall of the box at $x = a$ is *suddenly* moved to $x = 10a$, calculate the probability of finding the particle in

(a) the fourth excited ($n = 5$) state of the new box and

(b) the ninth ($n = 10$) excited state of the new box.

Exercise 10.15

A particle of mass m in the ground state of a one-dimensional harmonic oscillator is placed in a perturbation $\hat{V}(t) = -V_0\hat{x}e^{-t/\tau}$. Calculate to first-order perturbation theory the probability of finding the particle in its first excited state after a long time.

Exercise 10.16

A particle, initially (i.e., when $t < 0$) in its first excited state in an infinite potential well whose walls are located at $x = 0$ and $x = a$, is subject, starting at time $t = 0$, to a time-dependent perturbation $\hat{V}(t) = V_0\tau\hat{x}/(t^2 + \tau^2)$ where V_0 is a small real number. Calculate the probability that the particle will be found in its second excited state at $t = +\infty$.

Exercise 10.17

A one-dimensional harmonic oscillator has its spring constant suddenly reduced by half.

(a) If the oscillator is initially in its ground state, find the probability that the oscillator remains in the ground state.

(b) Find the work associated with this process.

Exercise 10.18

(a) Find the total transition rate associated with the decay of a harmonic oscillator, of charge q and mass m, from the nth excited state to the state just below.

(b) Find the power radiated by this oscillator as a result of its decay.

(c) Find the lifetime of the nth excited state.

(d) Estimate the order of magnitudes for the transition rate, the power, and the lifetime of the fifth excited state ($n = 5$) in the case of a harmonically oscillating electron (i.e., $q = e$) for the case of an optical radiation $\omega \simeq 10^{15}\,\text{rad}\,\text{s}^{-1}$.

Exercise 10.19

Assuming that $\langle\psi_f \mid \vec{r} \mid \psi_i\rangle$ is roughly equal to the size of the system under study, use a crude calculation to estimate the mean lifetime of

(a) an electric dipole transition in an atom where $\hbar\omega \sim 10\,\text{eV}$ and

(b) an electric dipole transition in a nucleus where $\hbar\omega \sim 1\,\text{MeV}$.

Exercise 10.20

A particle is initially (i.e., when $t < 0$) in its ground state in the potential $V(x) = -V_0\delta(x)$ with $V_0 > 0$.

(a) If the strength of the potential is changed *slowly* to $3V_0$, find the energy and wave function of the particle in the new potential.

(b) Calculate the work done with this process. Find a numerical value for this work in MeV if this particle were an electron and $V_0 = 200\,\text{MeV fm}$.

(c) If the strength of the potential is changed *suddenly* to $3V_0$, calculate the probability of finding the particle in the ground state of the new potential.

Exercise 10.21
A hydrogen atom in its ground state is placed at time $t = 0$ in a uniform electric field in the y-direction, $\vec{E}(t) = E_0 \hat{j} e^{-t^2/\tau^2}$. Calculate to first-order perturbation theory the probability that the atom will be found in any of the $n = 2$ states after a sufficiently long time ($t = +\infty$).

Exercise 10.22
A particle, initially (i.e., when $t < 0$) in its ground state in an infinite potential well with its walls at $x = 0$ and $x = a$, is subject, starting at time $t = 0$, to a time-dependent perturbation $\hat{V}(t) = V_0 \hat{x} \delta(x - 3a/4) e^{-t/\tau}$ where V_0 is a small parameter. Calculate the probability that the particle will be found in its first excited state ($n = 2$) at $t = +\infty$.

Exercise 10.23
Consider an isotropic (three-dimensional) harmonic oscillator which undergoes a transition from the second to the first excited state (i.e., 2s $\rightarrow$ 1p).
(a) Calculate the transition rate corresponding to 2s $\rightarrow$ 1p.
(b) Find the intensity associate with the 2s $\rightarrow$ 1p transition.

Exercise 10.24
Consider a particle which is initially (i.e., when $t < 0$) in its ground state in a three-dimensional box potential

$$V(x, y, z) = \begin{cases} 0, & 0 < x < a,\ 0 < y < 2a,\ 0 < z < 4a, \\ +\infty, & \text{elsewhere.} \end{cases}$$

(a) Find the energies and wave functions of the ground state and first excited state.
(b) This particle is then subject, starting at time $t = 0$, to a time-dependent perturbation $\hat{V}(t) = V_0 \hat{x}\hat{z} e^{-t^2}$ where V_0 is a small parameter. Calculate the probability that the particle will be found in the first excited state after a long time $t = +\infty$.

Chapter 11

Scattering Theory

Much of our understanding about the structure of matter is extracted from the scattering of particles. Had it not been for scattering, the structure of the microphysical world would have remained inaccessible to humans. It is through scattering experiments that important building blocks of matter, such as the atomic nucleus, the nucleons, and the various quarks, have been discovered.

11.1 Scattering and Cross Section

In a scattering experiment, one observes the collisions between a beam of incident particles and a target material. The total number of collisions over the duration of the experiment is proportional to the total number of incident particles and to the number of target particles per unit area in the path of the beam. In these experiments, one counts the collision products that come out of the target. After scattering, those particles that do not interact with the target continue their motion (undisturbed) in the forward direction, but those that interact with the target get scattered (deflected) at some angle as depicted in Figure 11.1. The number of particles coming out varies from one direction to the other. The number of particles scattered into an element of solid angle $d\Omega$ ($d\Omega = \sin\theta\, d\theta\, d\varphi$) is proportional to a quantity that plays a central role in the physics of scattering: the *differential cross section*. The differential cross section, which is denoted by $d\sigma(\theta,\varphi)/d\Omega$, is defined as the number of particles scattered into an element of solid angle $d\Omega$ in the direction (θ,φ) per unit time and incident flux:

$$\frac{d\sigma(\theta,\varphi)}{d\Omega} = \frac{1}{J_{inc}}\frac{dN(\theta,\varphi)}{d\Omega}, \tag{11.1}$$

where J_{inc} is the incident flux (or incident current density); it is equal to the number of incident particles per area per unit time. We can verify that $d\sigma/d\Omega$ has the dimensions of an area; hence it is appropriate to call it a differential cross section.

The relationship between $d\sigma/d\Omega$ and the total cross section σ is obvious:

$$\sigma = \int \frac{d\sigma}{d\Omega}d\Omega = \int_0^{\pi} \sin\theta\, d\theta \int_0^{2\pi} \frac{d\sigma(\theta,\varphi)}{d\Omega}d\varphi. \tag{11.2}$$

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

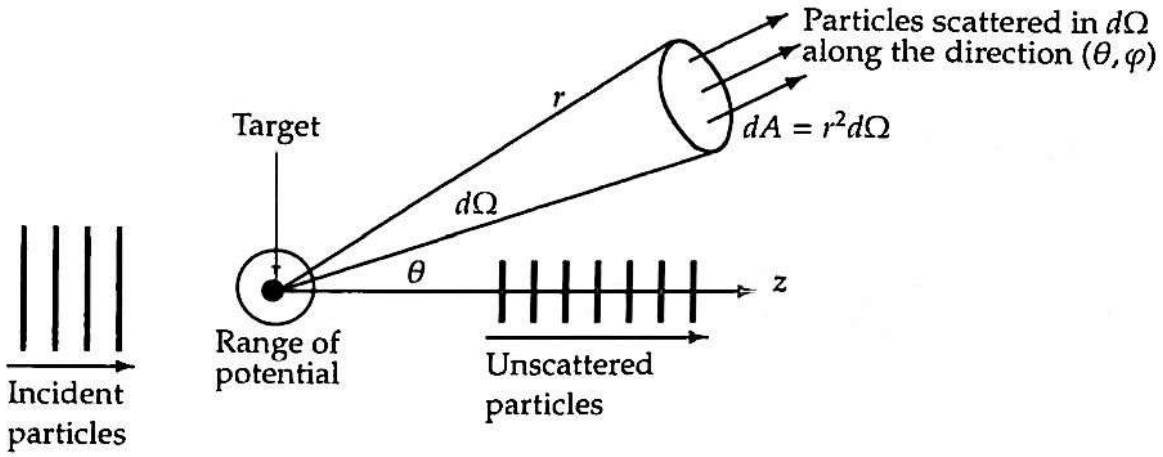


Figure 11.1: Scattering between an incident beam of particles and a fixed target: the scattered particles are detected within a solid angle $d\Omega$ along the direction (θ, φ).

Most scattering experiments are carried out in the laboratory (Lab) frame in which the target is initially at rest while the projectiles are moving. Calculations of the cross sections are generally easier to perform within the center of mass (CM) frame in which the center of mass of the projectiles–target system is at rest (before and after collision). In order to be able to compare the experimental measurements with the theoretical calculations, one has to know how to transform the cross sections from one frame into the other. We should note that the total cross section σ is the same in both frames, since the total number of collisions that take place does not depend on the frame in which the observation is carried out. As for the differential cross sections $d\sigma(\theta, \varphi)/d\Omega$, they are not the same in both frames, since the scattering angles (θ, φ) are frame dependent.

11.1.1 Connecting the Angles in the Lab and CM frames

To find the connection between the Lab and CM cross sections, we need first to find how the scattering angles in one frame are related to their counterparts in the other. Let us consider the scattering of two (structureless, nonrelativistic) particles of masses m_1 and m_2; m_2 represents the target, which is initially at rest, and m_1 the projectile. Figure 11.2 depicts such a scattering in the Lab and CM frames, where θ_1 and θ are the scattering angles of m_1 in the Lab and CM frames, respectively; we are interested in detecting m_1. In what follows we want to find the relation between θ_1 and θ. If $\vec{r}_{1_L}$ and $\vec{r}_{1_C}$ denote the position of m_1 in the Lab and CM frames, respectively, and if $\vec{R}$ denotes the position of the center of mass with respect to the Lab frame, we have $\vec{r}_{1_L} = \vec{r}_{1_C} + \vec{R}$. A time derivative of this relation leads to

$$\vec{V}_{1_L} = \vec{V}_{1_C} + \vec{V}_{CM}, \tag{11.3}$$

where $\vec{V}_{1_L}$ and $\vec{V}_{1_C}$ are the velocities of m_1 in the Lab and CM frames *before* collision and $\vec{V}_{CM}$ is the velocity of the CM with respect to the Lab frame. Similarly, the velocity of m_1 *after* collision is

$$\vec{V}'_{1_L} = \vec{V}'_{1_C} + \vec{V}_{CM}. \tag{11.4}$$

From Figure 11.2a we can infer the x and y components of (11.4):

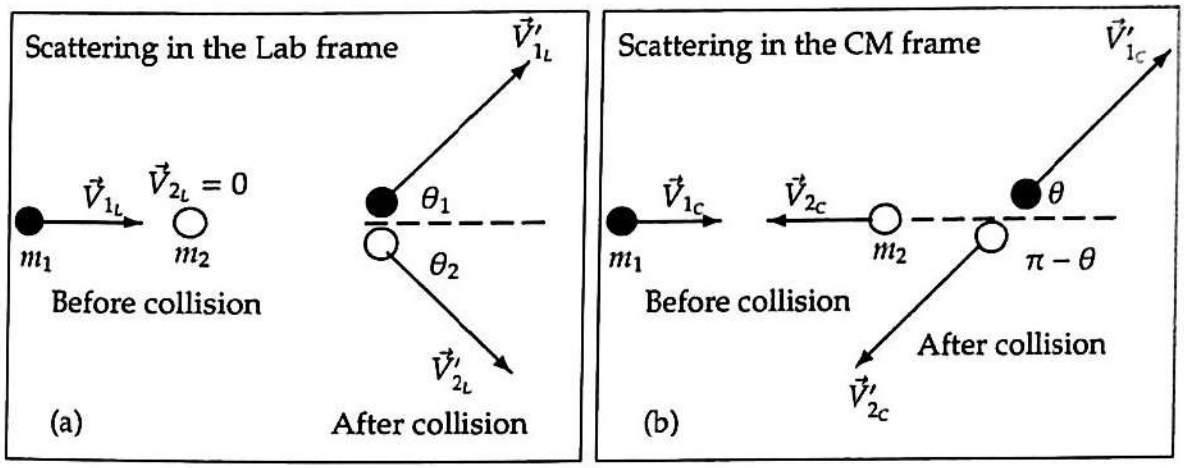


Figure 11.2: Elastic scattering of two structureless particles in the Lab and CM frames: a particle of mass m_1 strikes a particle m_2 initially at rest.

$$V'_{1_L} \cos\theta_1 = V'_{1_C} \cos\theta + V_{CM}, \tag{11.5}$$

$$V'_{1_L} \sin\theta_1 = V'_{1_C} \sin\theta. \tag{11.6}$$

Dividing (11.6) by (11.5), we end up with

$$\tan\theta_1 = \frac{\sin\theta}{\cos\theta + V_{CM}/V'_{1_C}}, \tag{11.7}$$

where V_{CM}/V'_{1_C} can be shown to be equal to m_1/m_2. To see this, since $\vec{V}_{2_L} = 0$, we have

$$\vec{V}_{CM} = \frac{m_1\vec{V}_{1_L} + m_2\vec{V}_{2_L}}{m_1 + m_2} = \frac{m_1}{m_1 + m_2}\vec{V}_{1_L}, \tag{11.8}$$

which when inserted into (11.3) leads to $\vec{V}_{1_L} = \vec{V}_{1_C} + m_1\vec{V}_{1_L}/(m_1 + m_2)$; hence

$$\vec{V}_{1_C} = \left(1 - \frac{m_1}{m_1 + m_2}\right)\vec{V}_{1_L} = \frac{m_2}{m_1 + m_2}\vec{V}_{1_L}. \tag{11.9}$$

On the other hand, since the center of mass is at rest in the CM frame, the total momenta before and after collisions are separately zero:

$$p_C = m_1 V_{1_C} - m_2 V_{2_C} = 0 \implies V_{2_C} = \frac{m_1}{m_2} V_{1_C}, \tag{11.10}$$

$$p'_{C_x} = m_1 V'_{1_C} \cos\theta - m_2 V'_{2_C} \cos\theta = 0 \implies V'_{2_C} = \frac{m_1}{m_2} V'_{1_C}. \tag{11.11}$$

In the case of *elastic* collision, the speeds of the particles in the CM frame are the same before and after collision; to see this, since the kinetic energy is conserved, a substitution of (11.10) and (11.11) into $\frac{1}{2}m_1V_{1_C}^2 + \frac{1}{2}m_2V_{2_C}^2 = \frac{1}{2}m_1V'^2_{1_C} + \frac{1}{2}m_2V'^2_{2_C}$ yields $V'_{1_C} = V_{1_C}$ and $V'_{2_C} = V_{2_C}$. Thus, we can rewrite (11.9) as

$$\vec{V}'_{1_C} = \vec{V}_{1_C} = \frac{m_2}{m_1 + m_2}\vec{V}_{1_L}. \tag{11.12}$$

Dividing (11.8) by (11.12) we obtain

$$\frac{V_{CM}}{V'_{1c}} = \frac{m_1}{m_2}. \tag{11.13}$$

Finally, a substitution of (11.13) into (11.7) yields

$$\tan\theta_1 = \frac{\sin\theta}{\cos\theta + V_{2c}/V_{1c}} = \frac{\sin\theta}{\cos\theta + m_1/m_2}, \tag{11.14}$$

which, using $\cos\theta_1 = 1/\sqrt{\tan^2\theta_1 + 1}$, becomes

$$\boxed{\cos\theta_1 = \frac{\cos\theta + \frac{m_1}{m_2}}{\sqrt{1 + \frac{m_1^2}{m_2^2} + 2\frac{m_1}{m_2}\cos\theta}}.} \tag{11.15}$$

Remark

By analogy with the foregoing analysis, we can establish a connection between θ_2 and θ. From (11.4) we have $\vec{V}'_{2_L} = \vec{V}'_{2c} + \vec{V}_{CM}$. The x and y components of this relation are

$$V'_{2_L}\cos\theta_2 = -V'_{2c}\cos\theta + V_{CM} = (-\cos\theta + 1)V'_{2c}, \tag{11.16}$$

$$V'_{2_L}\sin\theta_2 = -V'_{2c}\sin\theta; \tag{11.17}$$

in deriving (11.16), we have used $V_{CM} = V'_{2c} = V_{2c}$. A combination of (11.16) and (11.17) leads to

$$\boxed{\tan\theta_2 = \frac{\sin\theta}{-\cos\theta + V_{CM}/V'_{2c}} = \frac{\sin\theta}{1-\cos\theta} = \cot\left(\frac{\theta}{2}\right) \Longrightarrow \theta_2 = \frac{\pi - \theta}{2}.} \tag{11.18}$$

11.1.2 Connecting the Lab and CM Cross Sections

The connection between the differential cross sections in the Lab and CM frames can be obtained from the fact that the number of scattered particles passing through an infinitesimal cross section $d\sigma$ is the same in both frames: $d\sigma(\theta_1, \varphi_1) = d\sigma(\theta, \varphi)$. What differs is the solid angle $d\Omega$, since it is given in the Lab frame by $d\Omega_1 = \sin\theta_1 d\theta_1 d\varphi_1$ and in the CM frame by $d\Omega = \sin\theta\, d\theta d\varphi$. Thus, we have

$$\left(\frac{d\sigma}{d\Omega_1}\right)_{Lab} d\Omega_1 = \left(\frac{d\sigma}{d\Omega}\right)_{CM} d\Omega \Longrightarrow \left(\frac{d\sigma}{d\Omega_1}\right)_{Lab} = \left(\frac{d\sigma}{d\Omega}\right)_{CM} \frac{\sin\theta}{\sin\theta_1}\frac{d\theta}{d\theta_1}\frac{d\varphi}{d\varphi_1}, \tag{11.19}$$

where (θ_1, φ_1) are the scattering angles of particle m_1 in the Lab frame and (θ, φ) are its angles in the CM frame. Since there is cylindrical symmetry around the direction of the incident beam, we have $\varphi = \varphi_1$ and hence

$$\left(\frac{d\sigma}{d\Omega_1}\right)_{Lab} = \left(\frac{d\sigma}{d\Omega}\right)_{CM} \frac{d(\cos\theta)}{d(\cos\theta_1)}. \tag{11.20}$$

From (11.15) we have

$$\frac{d\cos\theta_1}{d\cos\theta} = \frac{1 + \frac{m_1}{m_2}\cos\theta}{\left(1 + \frac{m_1^2}{m_2^2} + 2\frac{m_1}{m_2}\cos\theta\right)^{3/2}}, \tag{11.21}$$

which when substituted into (11.20) leads to

$$\boxed{\left(\frac{d\sigma}{d\Omega_1}\right)_{Lab} = \frac{(1 + \frac{m_1^2}{m_2^2} + 2\frac{m_1}{m_2}\cos\theta)^{3/2}}{1 + \frac{m_1}{m_2}\cos\theta}\left(\frac{d\sigma}{d\Omega}\right)_{CM}.} \tag{11.22}$$

Similarly, we can show that (11.20) and (11.18) yield

$$\boxed{\left(\frac{d\sigma}{d\Omega_2}\right)_{Lab} = 4\cos\theta_2\left(\frac{d\sigma}{d\Omega_2}\right)_{CM} = 4\sin\left(\frac{\theta}{2}\right)\left(\frac{d\sigma}{d\Omega_2}\right)_{CM}.} \tag{11.23}$$

Limiting cases: (a) If $m_2 \gg m_1$, or when $\frac{m_1}{m_2} \to 0$, the Lab and CM results are the same, since (11.15) leads to $\theta_1 = \theta$ and (11.22) to $\left(\frac{d\sigma}{d\Omega_1}\right)_{Lab} = \left(\frac{d\sigma}{d\Omega}\right)_{CM}$. (b) If $m_2 = m_1$ then (11.15) leads to $\tan\theta_1 = \tan(\theta/2)$ or to $\theta_1 = \theta/2$; in this case (11.22) yields $\left(\frac{d\sigma}{d\Omega_1}\right)_{Lab} = 4\left(\frac{d\sigma}{d\Omega}\right)_{CM}\cos(\theta/2)$.

Example 11.1
In an elastic collision between two particles of equal mass, show that the two particles come out at right angles with respect to each other in the Lab frame.

Solution
In the special case $m_1 = m_2$, equations (11.14) and (11.18) respectively become

$$\tan\theta_1 = \tan\left(\frac{\theta}{2}\right), \qquad \tan\theta_2 = \cot\left(\frac{\theta}{2}\right) = \tan\left(\frac{\pi}{2} - \frac{\theta}{2}\right). \tag{11.24}$$

These two equations yield

$$\theta_1 = \frac{\theta}{2}, \qquad \theta_2 = \frac{\pi}{2} - \frac{\theta}{2} = \frac{\pi}{2} - \theta_1; \tag{11.25}$$

hence $\theta_1 + \theta_2 = \pi/2$. In these cases, (11.22) and (11.23) yield

$$\left(\frac{d\sigma}{d\Omega_1}\right)_{Lab} = 4\left(\frac{d\sigma}{d\Omega}\right)_{CM}\cos\theta_1 = 4\left(\frac{d\sigma}{d\Omega}\right)_{CM}\cos\left(\frac{\theta}{2}\right), \tag{11.26}$$

$$\left(\frac{d\sigma}{d\Omega_2}\right)_{Lab} = 4\left(\frac{d\sigma}{d\Omega_2}\right)_{CM}\cos\theta_2 = 4\left(\frac{d\sigma}{d\Omega}\right)_{CM}\sin\left(\frac{\theta}{2}\right). \tag{11.27}$$

11.2 Scattering Amplitude of Spinless Particles

The foregoing discussion dealt with definitions of the cross section and how to transform it from the Lab to the CM frame; the conclusions reached apply to classical as well as to quantum mechanics. In this section we deal with the quantum description of scattering. For simplicity,

we consider the case of elastic[1] scattering between two spinless, nonrelativistic particles of masses m_1 and m_2. During the scattering process, the particles interact with one another. If the interaction is *time independent*, we can describe the two-particle system with stationary states

$$\Psi(\vec{r}_1, \vec{r}_2, t) = \psi(\vec{r}_1, \vec{r}_2)e^{-iE_T t/\hbar}, \tag{11.28}$$

where E_T is the total energy and $\psi(\vec{r}_1, \vec{r}_2)$ is a solution of the time-independent Schrödinger equation:

$$\left[-\frac{\hbar^2}{2m_1}\vec{\nabla}_1^2 - \frac{\hbar^2}{2m_2}\vec{\nabla}_2^2 + \hat{V}(\vec{r}_1, \vec{r}_2)\right]\psi(\vec{r}_1, \vec{r}_2) = E_T\psi(\vec{r}_1, \vec{r}_2); \tag{11.29}$$

$\hat{V}(\vec{r}_1, \vec{r}_2)$ is the potential representing the interaction between the two particles.

In the case where the interaction between m_1 and m_2 depends only on their relative distance $r = |\vec{r}_1 - \vec{r}_2|$ (i.e., $\hat{V}(\vec{r}_1, \vec{r}_2) = \hat{V}(r)$), we can, as seen in Chapter 6, reduce the eigenvalue problem (11.29) to two decoupled eigenvalue problems: one for the center of mass (CM), which moves like a free particle of mass $M = m_1 + m_2$ and which is of no concern to us here, and another for a fictitious particle with a reduced mass $\mu = m_1m_2/(m_1 + m_2)$ which moves in the potential $\hat{V}(r)$:

$$-\frac{\hbar^2}{2\mu}\vec{\nabla}^2\psi(\vec{r}) + \hat{V}(r)\psi(\vec{r}) = E\psi(\vec{r}). \tag{11.30}$$

The problem of scattering between two particles is thus reduced to solving this equation. We are going to show that the differential cross section in the CM frame can be obtained from an asymptotic form of the solution of (11.30). Its solutions can then be used to calculate the probability per unit solid angle per unit time that the particle μ is scattered into a solid angle $d\Omega$ in the direction (θ, φ); this probability is given by the differential cross section $d\sigma/d\Omega$. In quantum mechanics the incident particle is described by means of a wave packet that interacts with the target. The incident wave packet must be spatially large so that spreading during the experiment is not appreciable. It must be large compared to the target's size and yet small compared to the size of the Lab so that it does not overlap simultaneously with the target and detector. After scattering, the wave function consists of an unscattered part propagating in the forward direction and a scattered part that propagates along some direction (θ, φ).

We can view (11.30) as representing the scattering of a particle of mass μ from a fixed scattering center that is described by $V(r)$, where r is the distance from the particle μ to the center of $V(r)$. We assume that $V(r)$ has a *finite* range a. Thus the interaction between the particle and the potential occurs only in a limited region of space $r \leq a$, which is called the *range* of $V(r)$, or the scattering region. Outside the range, $r > a$, the potential vanishes, $V(r) = 0$; the eigenvalue problem (11.30) then becomes

$$\left(\nabla^2 + k_0^2\right)\phi_{inc}(\vec{r}) = 0, \tag{11.31}$$

where $k_0^2 = 2\mu E/\hbar^2$. In this case μ behaves as a *free* particle before collision and hence can be described by a *plane* wave

$$\phi_{inc}(\vec{r}) = Ae^{i\vec{k}_0\cdot\vec{r}}, \tag{11.32}$$

where $\vec{k}_0$ is the wave vector associated with the incident particle and A is a normalization factor. Thus, prior to the interaction with the target, the particles of the incident beam are independent of each other; they move like free particles, each with a momentum $\vec{p} = \hbar\vec{k}_0$.

[1] In *elastic* scattering, the internal states and the structure of the colliding particles do not change.

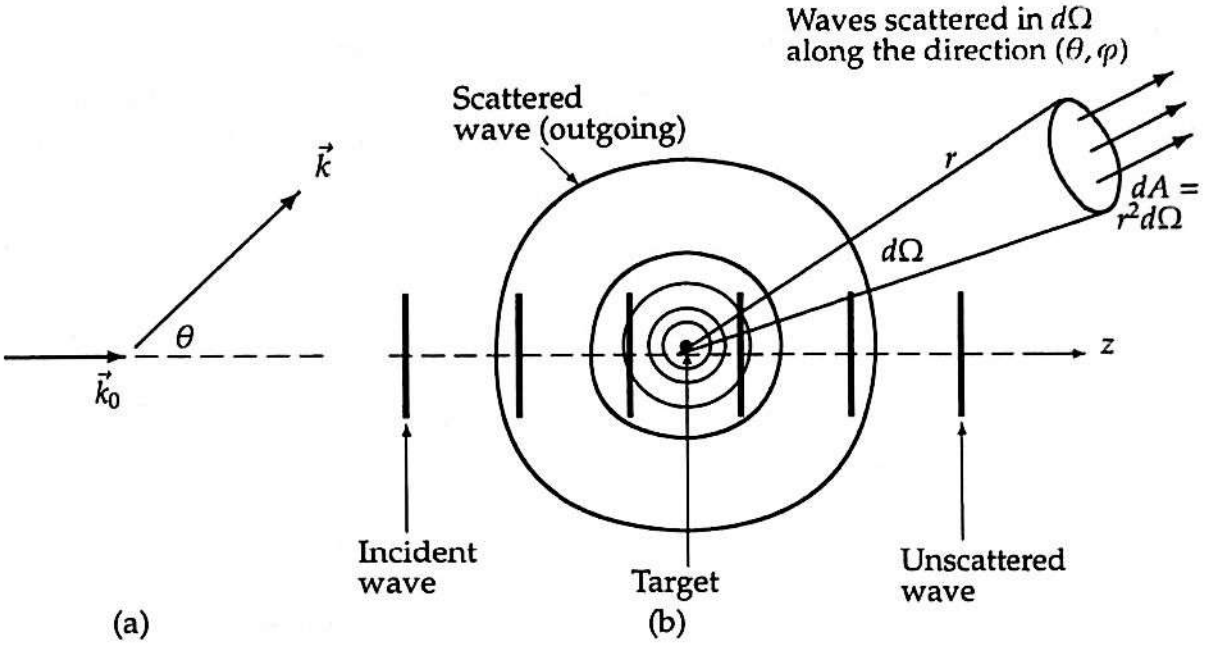


Figure 11.3: (a) Angle between the incident and scattered wave vectors $\vec{k}_0$ and $\vec{k}$. (b) Incident and scattered waves: the incident wave is a plane wave, $\phi_{inc}(\vec{r}) = Ae^{i\vec{k}_0\cdot\vec{r}}$, and the scattered wave, $\phi_{sc}(\vec{r}) = Af(\theta,\varphi)\frac{e^{ikr}}{r}$, is an outgoing wave.

When the incident wave (11.32) collides or interacts with the target, an outgoing wave $\phi_{sc}(\vec{r})$ is scattered out. In the case of an *isotropic* scattering, the scattered wave is *spherically symmetric*, having the form $e^{i\vec{k}\cdot\vec{r}}/r$. In general, however, the scattered wave is not spherically symmetric; its amplitude depends on the direction (θ,φ) along which it is detected and hence

$$\phi_{sc}(\vec{r}) = Af(\theta,\varphi)\frac{e^{i\vec{k}\cdot\vec{r}}}{r}, \tag{11.33}$$

where $f(\theta,\varphi)$ is called the *scattering amplitude*, $\vec{k}$ is the wave vector associated with the scattered particle, and θ is the angle between $\vec{k}_0$ and $\vec{k}$ as displayed in Figure 11.3a. After the scattering has taken place (Figure 11.3b), the total wave consists of a superposition of the incident plane wave (11.32) and the scattered wave (11.33):

$$\boxed{\psi(\vec{r}) = \phi_{inc}(\vec{r}) + \phi_{sc}(\vec{r}) \simeq A\left[e^{i\vec{k}_0\cdot\vec{r}} + f(\theta,\varphi)\frac{e^{i\vec{k}\cdot\vec{r}}}{r}\right]}, \tag{11.34}$$

where A is a normalization factor; since A has no effect on the cross section, as will be shown in (11.40), we will take it equal to one throughout the rest of the chapter. We now need to determine $f(\theta,\varphi)$ and $d\sigma/d\Omega$. In the following section we are going to show that the differential cross section is given in terms of the scattering amplitude by $d\sigma/d\Omega = |f(\theta,\varphi)|^2$.

11.2.1 Scattering Amplitude and Differential Cross Section

The scattering amplitude $f(\theta,\varphi)$ plays a central role in the theory of scattering, since it determines the differential cross section. To see this, let us first introduce the incident and scattered flux densities:

$$\vec{J}_{inc} = \frac{i\hbar}{2\mu}(\phi_{inc}\vec{\nabla}\phi^*_{inc} - \phi^*_{inc}\vec{\nabla}\phi_{inc}), \tag{11.35}$$

$$\vec{J}_{sc} = \frac{i\hbar}{2\mu}(\phi_{sc}\vec{\nabla}\phi^*_{sc} - \phi^*_{sc}\vec{\nabla}\phi_{sc}). \tag{11.36}$$

Inserting (11.32) into (11.35) and (11.33) into (11.36) and taking the magnitudes of the expressions thus obtained, we end up with

$$J_{inc} = |A|^2\frac{\hbar k_0}{\mu}, \qquad J_{sc} = |A|^2\frac{\hbar k}{\mu r^2}\left|f(\theta,\varphi)\right|^2. \tag{11.37}$$

Now, we may recall that the number $dN(\theta,\varphi)$ of particles scattered into an element of solid angle $d\Omega$ in the direction (θ,φ) and passing through a surface element $dA = r^2d\Omega$ per unit time is given as follows (see (11.1)):

$$dN(\theta,\varphi) = J_{sc}r^2d\Omega. \tag{11.38}$$

When combined with (11.37) this relation yields

$$\frac{dN}{d\Omega} = J_{sc}r^2 = |A|^2\frac{\hbar k}{\mu}\left|f(\theta,\varphi)\right|^2. \tag{11.39}$$

Now, inserting (11.39) and $J_{inc} = |A|^2\hbar k_0/\mu$ into (11.1), we end up with

$$\boxed{\frac{d\sigma}{d\Omega} = \frac{1}{J_{inc}}\frac{dN}{d\Omega} = \frac{k}{k_0}\left|f(\theta,\varphi)\right|^2.} \tag{11.40}$$

Since the normalization factor A does not contribute to the differential cross section, we will be taking it equal to one. For elastic scattering k_0 is equal to k; hence (11.40) reduces to

$$\boxed{\frac{d\sigma}{d\Omega} = \left|f(\theta,\varphi)\right|^2.} \tag{11.41}$$

The problem of determining the differential cross section $d\sigma/d\Omega$ therefore reduces to that of obtaining the scattering amplitude $f(\theta,\varphi)$.

11.2.2 Scattering Amplitude

We are going to show here that we can obtain the differential cross section in the CM frame from an asymptotic form of the solution of the Schrödinger equation (11.30). Let us first focus on the determination of $f(\theta,\varphi)$; it can be obtained from the solutions of (11.30), which in turn can be rewritten as

$$(\nabla^2 + k^2)\psi(\vec{r}) = \frac{2\mu}{\hbar^2}V(\vec{r})\psi(\vec{r}). \tag{11.42}$$

The general solution to this equation consists of a sum of two components: a general solution to the homogeneous equation:

$$(\nabla^2 + k_0^2)\psi_{homo}(\vec{r}) = 0, \tag{11.43}$$

and a particular solution to (11.42). First, note that $\psi_{homo}(\vec{r})$ is nothing but the incident plane wave (11.32). As for the particular solution to (11.42), we can express it in terms of *Green's function*. Thus, the general solution of (11.42) is given by

$$\psi(\vec{r}) = \phi_{inc}(\vec{r}) + \frac{2\mu}{\hbar^2}\int G(\vec{r}-\vec{r}\,')V(\vec{r}\,')\psi(\vec{r}\,')\,d^3r', \tag{11.44}$$

where $\phi_{inc}(\vec{r}) = e^{i\vec{k}_0\cdot\vec{r}}$ and $G(\vec{r}-\vec{r}\,')$ is Green's function corresponding to the operator on the left-hand side of (11.43). The function $G(\vec{r}-\vec{r}\,')$ is obtained by solving the point source equation

$$(\nabla^2 + k^2)G(\vec{r}-\vec{r}\,') = \delta(\vec{r}-\vec{r}\,'), \tag{11.45}$$

where $G(\vec{r}-\vec{r}\,')$ and $\delta(\vec{r}-\vec{r}\,')$ are given by their Fourier transforms as follows:

$$G(\vec{r}-\vec{r}\,') = \frac{1}{(2\pi)^3}\int e^{i\vec{q}\cdot(\vec{r}-\vec{r}\,')}\tilde{G}(\vec{q})\,d^3q, \tag{11.46}$$

$$\delta(\vec{r}-\vec{r}\,') = \frac{1}{(2\pi)^3}\int e^{i\vec{q}\cdot(\vec{r}-\vec{r}\,')}\,d^3q. \tag{11.47}$$

A substitution of (11.46) and (11.47) into (11.45) leads to

$$\left(-\vec{q}^{\,2} + \vec{k}^{\,2}\right)\tilde{G}(\vec{q}) = 1 \qquad \Longrightarrow \qquad \tilde{G}(\vec{q}) = \frac{1}{\vec{k}^{\,2} - \vec{q}^{\,2}}. \tag{11.48}$$

The expression for $G(\vec{r}-\vec{r}\,')$ is obtained by inserting (11.48) into (11.46):

$$G(\vec{r}-\vec{r}\,') = \frac{1}{(2\pi)^3}\int \frac{e^{i\vec{q}\cdot(\vec{r}-\vec{r}\,')}}{k^2-q^2}d^3q. \tag{11.49}$$

To integrate over the angles in

$$G(\vec{r}-\vec{r}\,') = \frac{1}{(2\pi)^3}\int_0^\infty \frac{q^2dq}{k^2-q^2}\int_0^\pi e^{iq|\vec{r}-\vec{r}\,'|\cos\theta}\sin\theta\,d\theta\int_0^{2\pi}d\varphi, \tag{11.50}$$

we need simply to make the variables change $x = \cos\theta$:

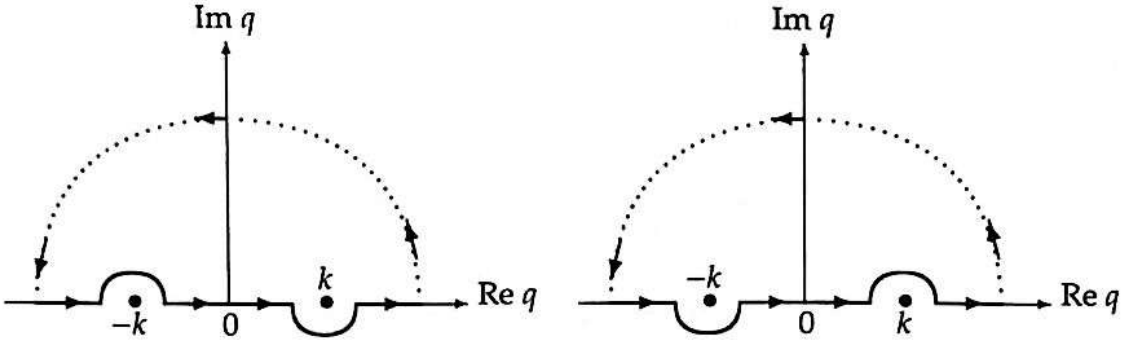
$$\int_0^\pi e^{iq|\vec{r}-\vec{r}\,'|\cos\theta}\sin\theta\,d\theta = \int_{-1}^1 e^{iq|\vec{r}-\vec{r}\,'|x}dx = \frac{1}{iq|\vec{r}-\vec{r}\,'|}\left(e^{iq|\vec{r}-\vec{r}\,'|} - e^{-iq|\vec{r}-\vec{r}\,'|}\right). \tag{11.51}$$

Hence (11.50) becomes

$$G(\vec{r}-\vec{r}\,') = \frac{1}{4\pi^2 i|\vec{r}-\vec{r}\,'|}\int_0^\infty \frac{q}{k^2-q^2}\left(e^{iq|\vec{r}-\vec{r}\,'|} - e^{-iq|\vec{r}-\vec{r}\,'|}\right)dq, \tag{11.52}$$

or

$$G(\vec{r}-\vec{r}\,') = -\frac{1}{4\pi^2 i|\vec{r}-\vec{r}\,'|}\int_{-\infty}^{+\infty}\frac{qe^{iq|\vec{r}-\vec{r}\,'|}}{q^2-k^2}dq. \tag{11.53}$$



(a) Contour for outgoing waves (b) Contour for incoming waves

Figure 11.4: Contours corresponding to outgoing and incoming waves.

We may evaluate this integral by the method of residues by closing the contour in the upper half of the q-plane: it is equal to $2\pi i$ times the residue of the integrand at the poles. Since there are two poles, $q = \pm k$, the integral has two possible values. The value corresponding to the pole at $q = k$, which lies inside the contour of integration in Figure 11.4a, is given by

$$G_+(\vec{r}-\vec{r}') = -\frac{1}{4\pi}\frac{e^{ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|} \tag{11.54}$$

and the value for the pole at $q = -k$ (Figure 11.4b) is

$$G_-(\vec{r}-\vec{r}') = -\frac{1}{4\pi}\frac{e^{-ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}. \tag{11.55}$$

Green's function $G_+(\vec{r}-\vec{r}')$ represents an *outgoing spherical wave* emitted from $\vec{r}'$ and the function $G_-(\vec{r}-\vec{r}')$ corresponds to an *incoming wave* that converges onto $\vec{r}'$. Since the scattered waves are outgoing waves, only $G_+(\vec{r}-\vec{r}')$ is of interest to us. Inserting (11.54) into (11.44) we obtain the total scattered wave function:

$$\boxed{\psi(\vec{r}) = \phi_{inc}(\vec{r}) - \frac{\mu}{2\pi\hbar^2}\int \frac{e^{ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}V(\vec{r}')\psi(\vec{r}')\,d^3r'.} \tag{11.56}$$

This is an *integral equation*; it does not yet give the unknown solution $\psi(\vec{r})$ but only contains it in the integrand. All we have done is to rewrite the Schrödinger (differential) equation (11.30) in an integral form (11.56), because the integral form is suitable for use in scattering theory. We are going to show that (11.56) reduces to (11.34) in the asymptotic limit $r \longrightarrow \infty$. But let us first mention that (11.56) can be solved approximately by means of a series of successive or iterative approximations, known as the *Born series*. The zero-order solution is given by $\psi_0(\vec{r}) = \phi_{inc}(\vec{r})$. The first-order solution $\psi_1(\vec{r})$ is obtained by inserting $\psi_0(\vec{r}) = \phi_{inc}(\vec{r})$ into the integral sign of (11.56):

$$\begin{aligned}\psi_1(\vec{r}) &= \phi_{inc}(\vec{r}) - \frac{\mu}{2\pi\hbar^2}\int \frac{e^{ik|\vec{r}-\vec{r}_1|}}{|\vec{r}-\vec{r}_1|}V(\vec{r}_1)\psi_0(\vec{r}_1)\,d^3r_1 \\ &= \phi_{inc}(\vec{r}) - \frac{\mu}{2\pi\hbar^2}\int \frac{e^{ik|\vec{r}-\vec{r}_1|}}{|\vec{r}-\vec{r}_1|}V(\vec{r}_1)\phi_{inc}(\vec{r}_1)\,d^3r_1.\end{aligned} \tag{11.57}$$

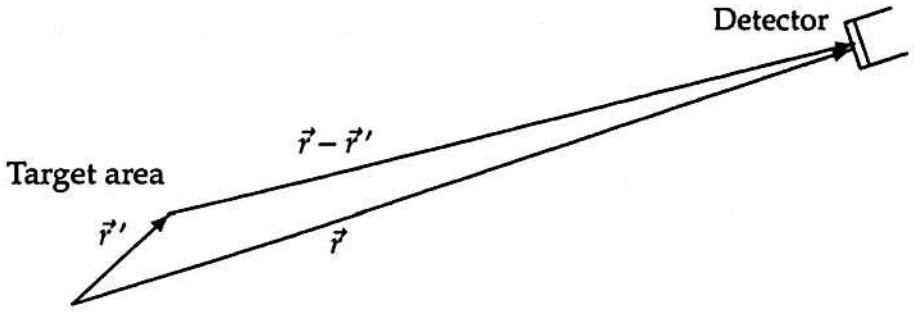


Figure 11.5: The distance r from the target to the detector is too large compared to the size r' of the target: $r \gg r'$.

The second order is obtained by inserting $\psi_1(\vec{r})$ into (11.56):

$$\begin{aligned}\psi_2(\vec{r}) &= \phi_{inc}(\vec{r}) - \frac{\mu}{2\pi\hbar^2}\int \frac{e^{ik|\vec{r}-\vec{r}_2|}}{|\vec{r}-\vec{r}_2|}V(\vec{r}_2)\psi_1(\vec{r}_2)\,d^3r_2 \\ &= \phi_{inc}(\vec{r}) - \frac{\mu}{2\pi\hbar^2}\int \frac{e^{ik|\vec{r}-\vec{r}_2|}}{|\vec{r}-\vec{r}_2|}V(\vec{r}_2)\phi_{inc}(\vec{r}_2)\,d^3r_2 \\ &\quad + \left(\frac{\mu}{2\pi\hbar^2}\right)^2 \int \frac{e^{ik|\vec{r}-\vec{r}_2|}}{|\vec{r}-\vec{r}_2|}V(\vec{r}_2)\,d^3r_2 \int \frac{e^{ik|\vec{r}_2-\vec{r}_1|}}{|\vec{r}_2-\vec{r}_1|}V(\vec{r}_1)\phi_{inc}(\vec{r}_1)\,d^3r_1.\end{aligned} \tag{11.58}$$

Continuing in this way, we can obtain $\psi(\vec{r})$ to the desired order; the nth order approximation for the wave function is a series which can be obtained by analogy with (11.57) and (11.58).

Asymptotic limit of the wave function

We are now going to show that (11.56) reduces to (11.34) for large values of r. In a scattering experiment, since the detector is located at distances (away from the target) that are much larger than the size of the target (Figure 11.5), we have $r \gg r'$, where r represents the distance from the target to the detector and r' the size of the detector. If $r \gg r'$ we may approximate $k|\vec{r}-\vec{r}'|$ and $|\vec{r}-\vec{r}'|^{-1}$ by

$$k|\vec{r}-\vec{r}'| = k\sqrt{\vec{r}^2 - 2\vec{r}\cdot\vec{r}' + \vec{r}'^2} \simeq kr - k\frac{\vec{r}}{r}\cdot\vec{r}' = kr - \vec{k}\cdot\vec{r}', \tag{11.59}$$

$$\frac{1}{|\vec{r}-\vec{r}'|} = \frac{1}{r}\frac{1}{\left|1-\vec{r}\cdot\vec{r}'/\vec{r}^2\right|} \simeq \frac{1}{r}\left(1+\frac{\vec{r}\cdot\vec{r}'}{r^2}\right) \simeq \frac{1}{r}, \tag{11.60}$$

where $\vec{k} = k\hat{r}$ is the wave vector associated with the scattered particle. From the previous two approximations, we may write the asymptotic form of (11.56) as follows:

$$\psi(\vec{r}) \longrightarrow e^{i\vec{k}_0\cdot\vec{r}} + \frac{e^{ikr}}{r}f(\theta,\varphi) \qquad (r\to\infty), \tag{11.61}$$

where

$$\boxed{f(\theta,\varphi) = -\frac{\mu}{2\pi\hbar^2}\int e^{-i\vec{k}\cdot\vec{r}'}V(\vec{r}')\psi(\vec{r}')\,d^3r' = -\frac{\mu}{2\pi\hbar^2}\langle\phi\mid\hat{V}\mid\psi\rangle,} \tag{11.62}$$

where $\phi(\vec{r})$ is a plane wave, $\phi(\vec{r}) = e^{i\vec{k}\cdot\vec{r}}$, and $\vec{k}$ is the wave vector of the scattered wave; the integration variable r' extends over the spatial degrees of freedom of the target. The differential cross section is then given by

$$\boxed{\frac{d\sigma}{d\Omega} = |f(\theta,\varphi)|^2 = \frac{\mu^2}{4\pi^2\hbar^4}\left|\int e^{-i\vec{k}\cdot\vec{r}'}\hat{V}(\vec{r}')\psi(\vec{r}')\,d^3r'\right|^2 = \frac{\mu^2}{4\pi^2\hbar^4}\left|\langle\phi\mid\hat{V}\mid\psi\rangle\right|^2.} \tag{11.63}$$

11.3 The Born Approximation

11.3.1 The First Born Approximation

If the potential $V(\vec{r})$ is weak enough, it will distort only slightly the incident plane wave. The *first Born approximation* consists then of approximating the scattered wave function $\psi(\vec{r})$ by a *plane wave*. This approximation corresponds to the first iteration of (11.56); that is, $\psi(\vec{r})$ is given by (11.57):

$$\psi(\vec{r}) \simeq \phi_{inc}(\vec{r}) - \frac{\mu i}{2\pi\hbar^2}\int \frac{e^{ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}V(\vec{r}')\phi_{inc}(\vec{r}')\,d^3r'. \tag{11.64}$$

Thus, using (11.62) and (11.63), we can write the scattering amplitude and the differential cross section in the first Born approximation as follows:

$$f(\theta,\varphi) = -\frac{\mu}{2\pi\hbar^2}\int e^{-i\vec{k}\cdot\vec{r}'}V(\vec{r}')\phi_{inc}(\vec{r}')\,d^3r' = -\frac{\mu}{2\pi\hbar^2}\int e^{i\vec{q}\cdot\vec{r}'}V(\vec{r}')\,d^3r', \tag{11.65}$$

$$\boxed{\frac{d\sigma}{d\Omega} = \left|f(\theta,\varphi)\right|^2 = \frac{\mu^2}{4\pi^2\hbar^4}\left|\int e^{i\vec{q}\cdot\vec{r}'}V(\vec{r}')\,d^3r'\right|^2,} \tag{11.66}$$

where $\vec{q} = \vec{k}_0 - \vec{k}$ and $\hbar\vec{q}$ is the momentum transfer; $\hbar\vec{k}_0$ and $\hbar\vec{k}$ are the linear momenta of the incident and scattered particles, respectively.

In *elastic scattering*, the magnitudes of $\vec{k}_0$ and $\vec{k}$ are equal (Figure 11.6); hence

$$q = \left|\vec{k}_0 - \vec{k}\right| = \sqrt{k_0^2 + k^2 - 2kk_0\cos\theta} = k\sqrt{2(1-\cos^2\theta)} = 2k\sin\left(\frac{\theta}{2}\right). \tag{11.67}$$

If the potential $V(\vec{r}')$ is *spherically symmetric*, $V(\vec{r}') = V(r')$, and choosing the z-axis along $\vec{q}$ (Figure 11.6), then $\vec{q}\cdot\vec{r}' = qr'\cos\theta'$ and therefore

$$\begin{aligned}\int e^{i\vec{q}\cdot\vec{r}'}V(\vec{r}')\,d^3r' &= \int_0^\infty r'^2V(r')\,dr'\int_0^\pi e^{iqr'\cos\theta'}\sin\theta'd\theta'\int_0^{2\pi}d\varphi' \\ &= 2\pi\int_0^\infty r'^2V(\vec{r}')\,dr'\int_{-1}^1 e^{iqr'x}dx = \frac{4\pi}{q}\int_0^\infty r'V(r')\sin(qr')\,dr'.\end{aligned} \tag{11.68}$$

Inserting (11.68) into (11.65) and (11.66) we obtain

$$\boxed{f(\theta) = -\frac{2\mu}{\hbar^2 q}\int_0^\infty r'V(r')\sin(qr')\,dr',} \tag{11.69}$$

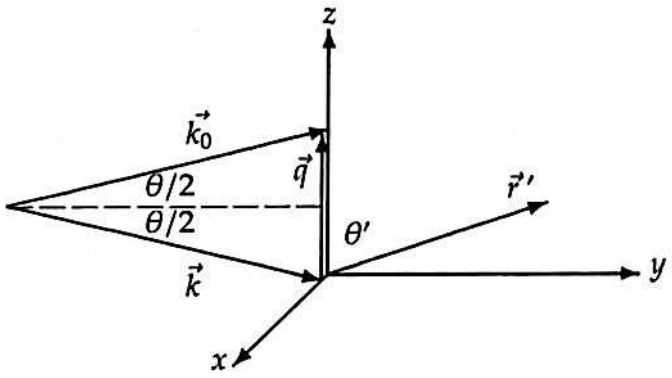


Figure 11.6: Momentum transfer for elastic scattering: $q = |\vec{k_0} - \vec{k}| = 2k\sin(\theta/2)$, $k_0 = k$.

$$\boxed{\frac{d\sigma}{d\Omega} = \left|f(\theta)\right|^2 = \frac{4\mu^2}{\hbar^4 q^2}\left|\int_0^\infty r' V(r')\sin(qr')\,dr'\right|^2.} \tag{11.70}$$

In summary, we have shown that by solving the Schrödinger equation (11.30) to first-order Born approximation (where the potential $V(\vec{r})$ is weak enough that the scattered wave function is only slightly different from the incident plane wave), the differential cross section is given by equation (11.70) for a spherically symmetric potential.

11.3.2 Validity of the First Born Approximation

The first Born approximation is valid whenever the wave function $\psi(\vec{r})$ is only slightly different from the incident plane wave; that is, whenever the second term in (11.64) is very small compared to the first:

$$\left|\frac{\mu}{2\pi\hbar^2}\int \frac{e^{ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}V(r')e^{i\vec{k_0}\cdot\vec{r}'}\,d^3r'\right| \ll \left|\phi_{inc}(\vec{r})\right|^2. \tag{11.71}$$

Since $\phi_{inc} = e^{i\vec{k_0}\cdot\vec{r}}$ we have

$$\left|\frac{\mu}{2\pi\hbar^2}\int \frac{e^{ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}V(r')e^{i\vec{k_0}\cdot\vec{r}'}\,d^3r'\right| \ll 1. \tag{11.72}$$

In elastic scattering $k_0 = k$ and assuming that the scattering potential is largest near $r = 0$, we have

$$\left|\frac{\mu}{\hbar^2}\int_0^\infty r' e^{ikr'}V(r')\,dr'\int_0^\pi e^{ikr'\cos\theta'}\sin\theta'\,d\theta'\right| \ll 1 \tag{11.73}$$

or

$$\boxed{\frac{\mu}{\hbar^2 k}\left|\int_0^\infty V(r')\left(e^{2ikr'}-1\right)dr'\right| \ll 1.} \tag{11.74}$$

Since the energy of the incident particle is proportional to k (it is purely kinetic, $E_i = \hbar^2k^2/2\mu$), we infer from (11.74) that the Born approximation is valid for large incident energies and weak scattering potentials. That is, when the average interaction energy between the incident

particle and the scattering potential is much smaller than the particle's incident kinetic energy, the scattered wave can be considered to be a plane wave.

Example 11.2

(a) Calculate the differential cross section in the first Born approximation for a Coulomb potential $V(r) = Z_1Z_2e^2/r$, where Z_1e and Z_2e are the charges of the projectile and target particles, respectively.

(b) To have a quantitative idea about the cross section derived in (a), consider the scattering of an alpha particle (i.e., a helium nucleus with $Z_1 = 2$ and $A_1 = 4$) from a gold nucleus ($Z_2 = 79$ and $A_2 = 197$). (i) If the scattering angle of the alpha particle in the Lab frame is $\theta_1 = 60°$, find its scattering angle θ in the CM frame. (ii) If the incident energy of the alpha particle is 8 MeV, find a numerical estimate for the cross section derived in (a).

Solution

(a) In the case of a Coulomb potential, $V(r) = Z_1Z_2e^2/r$, equation (11.70) becomes

$$\frac{d\sigma}{d\Omega} = \frac{4Z_1^2Z_2^2e^4\mu^2}{\hbar^4q^2}\left|\int_0^\infty \sin(qr)\,dr\right|^2, \tag{11.75}$$

where

$$\begin{aligned}\int_0^\infty \sin(qr)\,dr &= \lim_{\lambda\to0}\int_0^\infty e^{-\lambda r}\sin(qr)\,dr = \frac{1}{2i}\lim_{\lambda\to0}\left[\int_0^\infty e^{-(\lambda-iq)r}\,dr - \int_0^\infty e^{-(\lambda+iq)r}\,dr\right] \\ &= \frac{1}{2i}\lim_{\lambda\to0}\left[\frac{1}{\lambda-iq} - \frac{1}{\lambda+iq}\right] = \frac{1}{q}.\end{aligned} \tag{11.76}$$

Now, since $q = 2k\sin(\theta/2)$, an insertion of (11.76) into (11.75) leads to

$$\frac{d\sigma}{d\Omega} = \left(\frac{2Z_1\mu Z_2e^2}{\hbar^2q^2}\right)^2 = \left(\frac{Z_1Z_2\mu e^2}{2\hbar^2k^2}\right)^2\sin^{-4}\left(\frac{\theta}{2}\right) = \frac{Z_1^2Z_2^2e^4}{16E^2}\sin^{-4}\left(\frac{\theta}{2}\right), \tag{11.77}$$

where $E = \hbar^2k^2/2\mu$ is the kinetic energy of the incident particle. This relation is known as the *Rutherford formula* or the Coulomb cross section.

(b) (i) Since the mass ratio of the alpha particle to the gold nucleus is roughly equal to the ratio of their atomic masses, $m_1/m_2 = A_1/A_2 = \frac{4}{197} = 0.0203$, and since $\theta_1 = 60°$, equation (11.14) yields the value of the scattering angle in the CM frame:

$$\tan 60° = \frac{\sin\theta}{\cos\theta + 0.0203} \qquad \Longrightarrow \qquad \theta = 61°. \tag{11.78}$$

(ii) The numerical estimate of the cross section can be made easier by rewriting (11.77) in terms of the fine structure constant $\alpha = e^2/\hbar c = \frac{1}{137}$ and $\hbar c = 197.33$ MeV fm:

$$\frac{d\sigma}{d\Omega} = \frac{Z_1^2Z_2^2}{16E^2}\left(\frac{e^2}{\hbar c}\right)^2(\hbar c)^2\sin^{-4}\left(\frac{\theta}{2}\right) = \left(\frac{Z_1Z_2\alpha}{4}\right)^2\left(\frac{\hbar c}{E}\right)^2\sin^{-4}\left(\frac{\theta}{2}\right). \tag{11.79}$$

Since $Z_1 = 2$, $Z_2 = 79$, $\theta = 61°$, $\alpha = \frac{1}{137}$, $\hbar c = 197.33\,\text{MeV fm}$, and $E = 8\,\text{MeV}$, we have

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \left(\frac{2\times 79}{4\times 137}\right)^2 \left(\frac{197.33\,\text{MeV fm}}{8\,\text{MeV}}\right)^2 \sin^{-4}(30.5°) \\ &= 762.23\,\text{fm}^2 = 7.62\times 10^{-28}\,\text{m}^2 = 7.62\,\text{barn},\end{aligned} \tag{11.80}$$

where $1\,\text{barn} = 10^{-28}\,\text{m}^2$.

11.4 Partial Wave Analysis

So far we have considered only an approximate calculation of the differential cross section where the interaction between the projectile particle and the scattering potential $V(\vec{r})$ is considered small compared with the energy of the incident particle. In this section we are going to calculate the cross section without placing any limitation on the strength of $V(\vec{r})$.

11.4.1 Partial Wave Analysis for Elastic Scattering

We assume here the potential to be *spherically symmetric*. The angular momentum of the incident particle will therefore be conserved; a particle scattering from a central potential will have the same angular momentum before and after collision. Assuming that the incident plane wave is in the z-direction and hence $\phi_{inc}(\vec{r}) = \exp(ikr\cos\theta)$, we may express it in terms of a superposition of angular momentum eigenstates, each with a definite angular momentum number l (Chapter 6):

$$e^{i\vec{k}\cdot\vec{r}} = e^{ikr\cos\theta} = \sum_{l=0}^{\infty} i^l(2l+1)j_l(kr)P_l(\cos\theta). \tag{11.81}$$

We can then examine how each of the partial waves is distorted by $V(r)$ after the particle scatters from the potential. The most general solution of the Schrödinger equation (11.30) is

$$\psi(\vec{r}) = \sum_{lm} C_{lm}R_{kl}(r)Y_{lm}(\theta,\varphi). \tag{11.82}$$

Since $V(r)$ is central, the system is symmetrical (rotationally invariant) about the z-axis. The scattered wave function must not then depend on the azimuthal angle φ; hence $m = 0$. Thus, as $Y_{l0}(\theta,\varphi) \sim P_l(\cos\theta)$, the scattered wave function (11.82) becomes

$$\psi(r,\theta) = \sum_{l=0}^{\infty} a_l R_{kl}(r)P_l(\cos\theta), \tag{11.83}$$

where $R_{kl}(r)$ obeys the following radial equation (Chapter 6):

$$\left[\frac{d^2}{dr^2} + k^2 - \frac{l(l+1)}{r^2}\right](rR_{kl}(r)) = \frac{2m}{\hbar^2}V(r)\,(rR_{kl}(r)). \tag{11.84}$$

Each term of (11.83), which is known as a *partial wave*, is a joint eigenfunction of $\vec{L}^2$ and $\hat{L}_z$. A substitution of (11.81) into (11.34) with $\varphi = 0$ gives

$$\psi(r,\theta) \simeq \sum_{l=0}^{\infty} i^l(2l+1)j_l(kr)P_l(\cos\theta) + f(\theta)\frac{e^{ikr}}{r}. \tag{11.85}$$

The scattered wave function is given, on the one hand, by (11.83) and, on the other hand, by (11.85).

In almost all scattering experiments, detectors are located at distances from the target that are much larger than the size of the target itself; thus, the measurements taken by detectors pertain to scattered wave functions at large values of r. In what follows we are going to show that, by establishing a connection between the asymptotic forms of (11.83) and (11.85), we can determine the scattering amplitude and hence the differential cross section.

First, since the limit of the Bessel function $j_l(kr)$ for large values of r (Chapter 6) is given by

$$j_l(kr) \longrightarrow \frac{\sin(kr - l\pi/2)}{kr} \qquad (r \longrightarrow \infty), \tag{11.86}$$

the asymptotic form of (11.85) is

$$\psi(r,\theta) \longrightarrow \sum_{l=0}^{\infty} i^l(2l+1)P_l(\cos\theta)\frac{\sin(kr - l\pi/2)}{kr} + f(\theta)\frac{e^{ikr}}{r}, \tag{11.87}$$

and since $\sin(kr - l\pi/2) = [(-i)^l e^{ikr} - i^l e^{-ikr}]/2i$, because $e^{\pm il\pi/2} = \left(e^{\pm i\pi/2}\right)^l = (\pm i)^l$, we can write (11.87) as

$$\psi(r,\theta) \longrightarrow -\frac{e^{-ikr}}{2ikr}\sum_{l=0}^{\infty} i^{2l}(2l+1)P_l(\cos\theta) + \frac{e^{ikr}}{r}\left[f(\theta) + \frac{1}{2ik}\sum_{l=0}^{\infty} i^l(-i)^l(2l+1)P_l(\cos\theta)\right]. \tag{11.88}$$

Second, to find the asymptotic form of (11.83), we need first to determine the asymptotic form of the radial function $R_{kl}(r)$. At large values of r, the scattering potential is effectively zero, for it is short range. In this case (11.84) becomes

$$\left(\frac{d^2}{dr^2} + k^2\right)(rR_{kl}(r)) = 0. \tag{11.89}$$

As seen in Chapter 6, the general solution of this equation is given by a linear combination of the spherical Bessel and Neumann functions

$$R_{kl}(r) = A_l j_l(kr) + B_l n_l(kr), \tag{11.90}$$

where the asymptotic form of the Neumann function is

$$n_l(kr) \longrightarrow -\frac{\cos(kr - l\pi/2)}{kr} \qquad (r \longrightarrow \infty). \tag{11.91}$$

Inserting of (11.86) and (11.91) into (11.90), we obtain the asymptotic form of the radial function:

$$R_{kl}(r) \longrightarrow A_l\frac{\sin(kr - l\pi/2)}{kr} - B_l\frac{\cos(kr - l\pi/2)}{kr} \qquad (r \longrightarrow \infty). \tag{11.92}$$

If $V(r) = 0$ for all r (free particle), the solution of (11.84), $rR_{kl}(r)$, must vanish at $r = 0$; thus $R_{kl}(r)$ must be finite at the origin. Since the Neumann function diverges at $r = 0$, the cosine term in (11.92) does not represent a physically acceptable solution; hence, it needs to be discarded near the origin. By rewriting (11.92) in the form

$$R_{kl}(r) \longrightarrow C_l \frac{\sin(kr - l\pi/2 + \delta_l)}{kr} \qquad (r \longrightarrow \infty), \tag{11.93}$$

we have $A_l = C_l \cos\delta_l$ and $B_l = -C_l \sin\delta_l$, hence $C_l = \sqrt{A_l^2 + B_l^2}$ and

$$\tan\delta_l = -\frac{B_l}{A_l} \qquad \Longrightarrow \qquad \delta_l = -\tan^{-1}\left(\frac{B_l}{A_l}\right). \tag{11.94}$$

We see that, with $\delta_l = 0$, the radial function $R_{kl}(r)$ of (11.93) is finite at $r = 0$, since (11.93) reduces to $j_l(kr)$. So δ_l is a real angle which vanishes for all values of l in the absence of the scattering potential (i.e., $V = 0$); δ_l is called the *phase shift*. It measures, at large values of r, the degree to which $R_{kl}(r)$ differs from $j_l(kr)$ (recall that $j_l(kr)$ is the radial function when there is no scattering). Since this "distortion," or the difference between $R_{kl}(r)$ and $j_l(kr)$, is due to the potential $V(r)$, we would expect the cross section to depend on δ_l. Using (11.93) we can write the asymptotic limit of (11.83) as

$$\psi(r,\theta) \longrightarrow \sum_{l=0}^{\infty} a_l P_l(\cos\theta) \frac{\sin(kr - l\pi/2 + \delta_l)}{kr} \qquad (r \longrightarrow \infty). \tag{11.95}$$

This wave function is known as a *distorted plane wave*, for it differs from a plane wave by having phase shifts δ_l. Since $\sin(kr - l\pi/2 + \delta_l) = [(-i)^l e^{ikr} e^{i\delta_l} - i^l e^{-ikr} e^{-i\delta_l}]/2i$, we can rewrite (11.95) as

$$\psi(r,\theta) \longrightarrow -\frac{e^{-ikr}}{2ikr} \sum_{l=0}^{\infty} a_l i^l e^{-i\delta_l} P_l(\cos\theta) + \frac{e^{ikr}}{2ikr} \sum_{l=0}^{\infty} a_l (-i)^l e^{i\delta_l} P_l(\cos\theta). \tag{11.96}$$

Up to now we have shown that the asymptotic forms of (11.83) and (11.85) are given by (11.96) and (11.88), respectively. Equating the coefficients of e^{-ikr}/r in (11.88) and (11.96), we obtain $(2l + 1)i^{2l} = a_l i^l e^{-i\delta_l}$ and hence

$$a_l = (2l + 1) i^l e^{i\delta_l}. \tag{11.97}$$

Substituting (11.97) into (11.96) and this time equating the coefficient of e^{ikr}/r in the resulting expression with that of (11.88), we have

$$f(\theta) + \frac{1}{2ik} \sum_{l=0}^{\infty} i^l (-i)^l (2l + 1) P_l(\cos\theta) = \frac{1}{2ik} \sum_{l=0}^{\infty} (2l + 1) i^l (-i)^l e^{2i\delta_l} P_l(\cos\theta), \tag{11.98}$$

which, when combined with $(e^{2i\delta_l} - 1)/2i = e^{i\delta_l} \sin\delta_l$ and $i^l(-i)^l = 1$, leads to

$$\boxed{f(\theta) = \sum_{l=0}^{\infty} f_l(\theta) = \frac{1}{2ik} \sum_{l=0}^{\infty} (2l + 1) P_l(\cos\theta)(e^{2i\delta_l} - 1) = \frac{1}{k} \sum_{l=0}^{\infty} (2l + 1) e^{i\delta_l} \sin\delta_l P_l(\cos\theta),} \tag{11.99}$$

where $f_l(\theta)$ is known as the partial wave amplitude.

From (11.99) we can obtain the differential and the total cross sections

$$\frac{d\sigma}{d\Omega} = \left| f(\theta) \right|^2 = \frac{1}{k^2} \sum_{l=0}^{\infty} \sum_{l'=0}^{\infty} (2l+1)(2l'+1) e^{i(\delta_l - \delta_{l'})} \sin\delta_l \sin\delta_{l'} P_l(\cos\theta) P_{l'}(\cos\theta), \tag{11.100}$$

$$\begin{aligned} \sigma &= \int \frac{d\sigma}{d\Omega} d\Omega = \int_0^{\pi} \left| f(\theta) \right|^2 \sin\theta \, d\theta \int_0^{2\pi} d\varphi = 2\pi \int_0^{\pi} \left| f(\theta) \right|^2 \sin\theta \, d\theta \\ &= \frac{2\pi}{k^2} \sum_{l=0}^{\infty} \sum_{l'=0}^{\infty} (2l+1)(2l'+1) e^{i(\delta_l - \delta_{l'})} \sin\delta_l \sin\delta_{l'} \int_0^{\pi} P_l(\cos\theta) P_{l'}(\cos\theta) \sin\theta \, d\theta. \end{aligned} \tag{11.101}$$

Using the relation $\int_0^{\pi} P_l(\cos\theta) P_{l'}(\cos\theta) \sin\theta \, d\theta = [2/(2l+1)]\delta_{ll'}$, we can reduce (11.101) to

$$\boxed{\sigma = \sum_{l=0}^{\infty} \sigma_l = \frac{4\pi}{k^2} \sum_{l=0}^{\infty} (2l+1) \sin^2\delta_l,} \tag{11.102}$$

where σ_l are called the *partial cross sections* corresponding to the scattering of particles in various angular momentum states. The differential cross section (11.100) consists of a superposition of terms with different angular momenta; this gives rise to *interference* patterns between different partial waves corresponding to different values of l. The interference terms go away in the total cross section when the integral over θ is carried out. Note that when $V = 0$ everywhere, all the phase shifts δ_l vanish, and hence the partial and total cross sections, as indicated by (11.100) and (11.102), are zero. Note that, as shown in equations (11.99) and (11.102), $f(\theta)$ and σ are given as infinite series over the angular momentum l. We may recall that, for cases of practical importance with the exception of the Coulomb potential, these series converge after a finite number of terms.

We should note that in the case where we have a scattering between particles that are in their respective s states, $l = 0$, the scattering amplitude (11.99) becomes

$$f_0 = \frac{1}{k} e^{i\delta_0} \sin\delta_0 \qquad (l = 0), \tag{11.103}$$

where we have used $P_0(\cos\theta) = 1$. Since f_0 does not depend on θ, the differential and total cross sections are given by the following simple relations:

$$\frac{d\sigma}{d\Omega} = \left| f_0 \right|^2 = \frac{1}{k^2} \sin^2\delta_0, \qquad \sigma = 4\pi \left| f_0 \right|^2 = \frac{4\pi}{k^2} \sin^2\delta_0 \qquad (l = 0). \tag{11.104}$$

An important issue here is the fact that the total cross section can be related to the *forward scattering amplitude* $f(0)$. Since $P_l(\cos\theta) = P_l(1) = 1$ when $\theta = 0$, equation (11.99) leads to

$$f(0) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) \left(\sin\delta_l \cos\delta_l + i \sin^2\delta_l \right), \tag{11.105}$$

which when combined with (11.102) yields the connection between $f(0)$ and σ:

$$\boxed{\frac{4\pi}{k}\,\mathrm{Im}\,f(0) = \sigma = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l.} \tag{11.106}$$

This is known as the *optical theorem* (it is reminiscent of a similar theorem in optics which deals with the scattering of light). The physical origin of this theorem is the conservation of particles (or probability). The beam emerging (after scattering) along the incidence direction ($\theta = 0$) contains fewer particles than the incident beam, since a number of particles have scattered in various directions. This decrease in the number of particles is measured by the total cross section σ; that is, the number of particles removed from the incident beam along the incidence direction is proportional to σ or, equivalently, to the imaginary part of $f(0)$. We should note that, although (11.106) was derived for elastic scattering, the optical theorem, as will be shown later, is also valid for inelastic scattering.

11.4.2 Partial Wave Analysis for Inelastic Scattering

The scattering amplitude (11.99) can be rewritten as

$$f(\theta) = \sum_{l=0}^{\infty}(2l+1)f_l(k)P_l(\cos\theta), \tag{11.107}$$

where

$$f_l(k) = \frac{1}{k}e^{i\delta_l}\sin\delta_l = \frac{1}{2ik}\left(e^{2i\delta_l}-1\right) = \frac{1}{2ik}\left(S_l(k)-1\right), \tag{11.108}$$

with

$$S_l(k) = e^{2i\delta_l}. \tag{11.109}$$

In the case where there is no flux loss, we must have $|S_l(k)| = 1$. However, this requirement is not valid whenever there is *absorption* of the incident beam. In this case of flux loss, $S_l(k)$ is redefined by

$$S_l(k) = \eta_l(k)e^{2i\delta_l}, \tag{11.110}$$

with $0 < \eta_l(k) \leq 1$; hence (11.108) and (11.107) become

$$f_l(k) = \frac{\eta_l e^{2i\delta_l}-1}{2ik} = \frac{1}{2k}\left[\eta_l\sin 2\delta_l + i(1-\eta_l\cos 2\delta_l)\right], \tag{11.111}$$

$$\boxed{f(\theta) = \frac{1}{2k}\sum_{l=0}^{\infty}(2l+1)\left[\eta_l\sin 2\delta_l + i(1-\eta_l\cos 2\delta_l)\right]P_l(\cos\theta).} \tag{11.112}$$

The *total elastic* scattering cross section is given by

$$\boxed{\sigma_{el} = 4\pi\sum_{l=0}^{\infty}(2l+1)|f_l|^2 = \frac{\pi}{k^2}\sum_{l}(2l+1)(1+\eta_l^2-2\eta_l\cos 2\delta_l).} \tag{11.113}$$

The *total inelastic* scattering cross section, which describes the loss of flux, is given by

$$\boxed{\sigma_{inel} = \frac{\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\left(1-\eta_l^2(k)\right).} \tag{11.114}$$

Thus, if $\eta_l(k) = 1$ there is no inelastic scattering, but if $\eta_l = 0$ we have total absorption, although there is still elastic scattering in this partial wave. The sum of (11.113) and (11.114) gives the total cross section:

$$\boxed{\sigma_{tot} = \sigma_{el} + \sigma_{inel} = \frac{2\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\left(1-\eta_l\cos(2\delta_l)\right).} \tag{11.115}$$

Next, using (11.107) and (11.111), we infer

$$\operatorname{Im} f(0) = \sum_{l=0}^{\infty}(2l+1)\operatorname{Im} f_l = \frac{1}{2k}\sum_{l=0}^{\infty}(2l+1)\left(1-\eta_l\cos(2\delta_l)\right). \tag{11.116}$$

A comparison of (11.115) and (11.116) gives the optical theorem relation, $\operatorname{Im} f(0) = k\sigma_{tot}/4\pi$; hence the optical theorem is also valid for inelastic scattering.

Example 11.3 (High-energy scattering from a black disk)
Discuss the scattering from a black disk at high energies.

Solution
A black disk is totally absorbing (i.e., $\eta_l(k) = 0$). Assuming the values of l do not exceed a maximum value l_{max} ($l \leq l_{max}$) and that k is large (high-energy scattering), we have $l_{max} = ka$ where a is the radius of the disk. Since $\eta_l = 0$, equations (11.113) and (11.114) lead to

$$\sigma_{inel} = \sigma_{el} = \frac{\pi}{k^2}\sum_{l=0}^{ka}(2l+1) = \frac{\pi}{k^2}(ka+1)^2 \simeq \pi a^2; \tag{11.117}$$

hence the total cross section is given by

$$\sigma_{tot} = \sigma_{el} + \sigma_{inel} = 2\pi a^2. \tag{11.118}$$

Classically, the total cross section of a disk is equal to πa^2. The factor 2 in (11.118) is due to purely quantum effects, since in the high-energy limit there are two kinds of scattering: one corresponding to waves that hit the disk, where the cross section is equal to the classical cross section πa^2, and the other to waves that are diffracted. According to Babinet's principle, the cross section for the waves diffracted by a disk is also equal to πa^2.

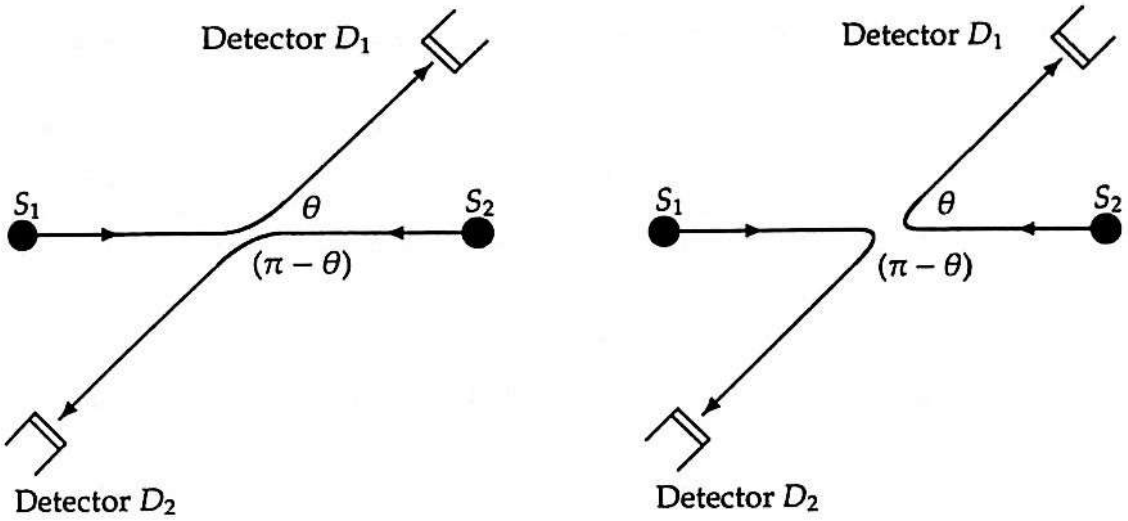


Figure 11.7: When scattering two *identical* particles in the center of mass frame, it is impossible to distinguish between the particle that scatters at angle θ from the one that scatters at $(\pi - \theta)$

11.5 Scattering of Identical Particles

First, let us consider the scattering of two identical *bosons* in their center of mass frame (we will consider the scattering of two identical fermions in a moment). *Classically*, the cross section for the scattering of two identical particles whose interaction potential is central is given by

$$\sigma_{cl}(\theta) = \sigma(\theta) + \sigma(\pi - \theta). \tag{11.119}$$

In quantum mechanics there is no way of distinguishing, as indicated in Figure 11.7, between the particle that scatters at an angle θ from the one that scatters at $(\pi - \theta)$. Thus, the scattered wave function must be symmetric:

$$\psi_{sym}(\vec{r}) \longrightarrow e^{i\vec{k_0}\cdot\vec{r}} + e^{-i\vec{k_0}\cdot\vec{r}} + f_{sym}(\theta)\frac{e^{ikr}}{r}, \tag{11.120}$$

and so must also be the scattering amplitude:

$$f_{boson}(\theta) = f(\theta) + f(\pi - \theta). \tag{11.121}$$

Therefore, the differential cross section is

$$\begin{aligned}\frac{d\sigma}{d\Omega}_{boson} &= \Big|f(\theta) + f(\pi - \theta)\Big|^2 = \Big|f(\theta)\Big|^2 + \Big|f(\pi - \theta)\Big|^2 + f(\theta)^* f(\pi - \theta) + f(\theta) f^*(\pi - \theta) \\ &= \Big|f(\theta)\Big|^2 + \Big|f(\pi - \theta)\Big|^2 + 2\,\mathrm{Re}\Big[f^*(\theta) f(\pi - \theta)\Big].\end{aligned} \tag{11.122}$$

In sharp contrast to its classical counterpart, equation (11.122) contains an interference term $2\,\mathrm{Re}\big[f^*(\theta)f(\pi - \theta)\big]$. Note that when $\theta = \pi/2$, we have $(d\sigma/d\Omega)_{boson} = 4\big|f(\pi/2)\big|^2$; this is twice as large as the classical expression (which has no interference term): $(d\sigma/d\Omega)_{cl} = 2\big|f(\pi/2)\big|^2$.

If the particles were distinguishable, the differential cross section will be four times smaller, $(d\sigma/d\Omega)_{distinguishable} = |f(\pi/2)|^2$.

Consider now the scattering of two identical spin $\frac{1}{2}$ particles. This is the case, for example, of electron–electron or proton–proton scattering. The wave function of a two spin $\frac{1}{2}$ particle system is known to be either symmetric or antisymmetric. When the spatial wave function is symmetric, that is the two particles are in a spin singlet state, the differential cross section is given by

$$\frac{d\sigma_S}{d\Omega} = |f(\theta) + f(\pi - \theta)|^2, \tag{11.123}$$

but when the two particles are in a spin triplet state, the spatial wave function is antisymmetric, and hence

$$\frac{d\sigma_A}{d\Omega} = |f(\theta) - f(\pi - \theta)|^2. \tag{11.124}$$

If the incident particles are unpolarized, the various spin states will be equally likely, so the triplet state will be three times as likely as the singlet:

$$\begin{aligned}\frac{d\sigma}{d\Omega}_{fermion} &= \frac{3}{4}\frac{d\sigma_a}{d\Omega} + \frac{1}{4}\frac{d\sigma_s}{d\Omega} = \frac{3}{4}|f(\theta) - f(\pi - \theta)|^2 + \frac{1}{4}|f(\theta) + f(\pi - \theta)|^2 \\ &= |f(\theta)|^2 + |f(\pi - \theta)|^2 - \mathrm{Re}\,[f^*(\theta)f(\pi - \theta)]. \end{aligned} \tag{11.125}$$

When $\theta = \pi/2$, we have $(d\sigma/d\Omega)_{fermion} = |f(\pi/2)|^2$; this quantum differential cross section is half the classical expression, $(d\sigma/d\Omega)_{cl} = 2|f(\pi/2)|^2$, and four times smaller than the expression corresponding to the scattering of two identical bosons, $(d\sigma/d\Omega)_{boson} = 4|f(\pi/2)|^2$.

We should note that, in the case of partial wave analysis for elastic scattering, using the relations $\cos(\pi - \theta) = -\cos\theta$ and $P_l(\cos(\pi - \theta)) = P_l(-\cos\theta) = (-1)^l P_l(\cos\theta)$ and inserting them into (11.99), we can write

$$\begin{aligned} f(\pi - \theta) &= \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)e^{i\delta_l}\sin\delta_l P_l(\cos(\pi - \theta)) \\ &= \frac{1}{k}\sum_{l=0}^{\infty}(-1)^l(2l+1)e^{i\delta_l}\sin\delta_l P_l(\cos\theta), \end{aligned} \tag{11.126}$$

and hence

$$f(\theta) \pm f(\pi - \theta) = \frac{1}{k}\sum_{l=0}^{\infty}\left[1 \pm (-1)^l\right](2l+1)e^{i\delta_l}\sin\delta_l P_l(\cos\theta). \tag{11.127}$$

Example 11.4

Calculate the differential cross section in the first Born approximation for the scattering between two identical particles having spin 1, mass m, and interacting through a potential $V(r) = V_0 e^{-ar}$.

Solution

As seen in Chapter 7, the spin states of two identical particles with spin $s_1 = s_2 = 1$ consist of a total of nine states: a quintuplet $|2, m\rangle$ (i.e., $|2, \pm2\rangle$, $|2, \pm1\rangle$, $|2, 0\rangle$) and a singlet $|0, 0\rangle$,

which are symmetric, and a triplet $|1, m\rangle$ (i.e., $|1, \pm 1\rangle$, $|1, 0\rangle$), which are antisymmetric under particle permutation. That is, while the six spin states corresponding to $S = 2$ and $S = 0$ are symmetric, the three $S = 1$ states are antisymmetric. Thus, if the scattering particles are unpolarized, the differential cross section is

$$\frac{d\sigma}{d\Omega} = \frac{5}{9}\frac{d\sigma_S}{d\Omega} + \frac{1}{9}\frac{d\sigma_S}{d\Omega} + \frac{3}{9}\frac{d\sigma_A}{d\Omega} = \frac{2}{3}\frac{d\sigma_S}{d\Omega} + \frac{1}{3}\frac{d\sigma_A}{d\Omega}, \tag{11.128}$$

where

$$\frac{d\sigma_S}{d\Omega} = \left| f(\theta) + f(\pi - \theta) \right|^2, \qquad \frac{d\sigma_A}{d\Omega} = \left| f(\theta) - f(\pi - \theta) \right|^2. \tag{11.129}$$

The scattering amplitude is given in the Born approximation by (11.69):

$$\begin{aligned} f(\theta) &= -\frac{2V_0\mu}{\hbar^2 q}\int_0^\infty re^{-ar}\sin(qr)\,dr = -\frac{V_0\mu}{i\hbar^2 q}\int_0^\infty re^{-(a-iq)r}\,dr + \frac{V_0\mu}{i\hbar^2 q}\int_0^\infty re^{-(a+iq)r}\,dr \\ &= \frac{V_0\mu}{\hbar^2 q}\frac{\partial}{\partial q}\int_0^\infty e^{-(a-iq)r}\,dr + \frac{V_0\mu}{\hbar^2 q}\frac{\partial}{\partial q}\int_0^\infty e^{-(a+iq)r}\,dr \\ &= \frac{V_0\mu}{\hbar^2 q}\frac{\partial}{\partial q}\left(\frac{1}{a-iq}\right) + \frac{V_0\mu}{\hbar^2 q}\frac{\partial}{\partial q}\left(\frac{1}{a+iq}\right) = \frac{V_0\mu}{\hbar^2 q}\left[\frac{i}{(a-iq)^2} + \frac{-i}{(a+iq)^2}\right] \\ &= -\frac{4V_0\mu a}{\hbar^2}\frac{1}{(a^2+q^2)^2} = -\frac{4V_0\mu a}{\hbar^2}\frac{1}{\left(a^2 + 4k^2\sin^2(\theta/2)\right)^2}, \end{aligned} \tag{11.130}$$

where we have used $q = 2k\sin(\theta/2)$, with $\mu = m/2$. Since $\sin[(\pi - \theta)/2] = \cos(\theta/2)$, we have

$$\frac{d\sigma_S}{d\Omega} = \frac{16V_0^2\mu^2a^2}{\hbar^4}\left[\frac{1}{\left(a^2 + 4k^2\sin^2(\theta/2)\right)^2} + \frac{1}{(a^2 + 4k^2\cos^2(\theta/2))^2}\right]^2, \tag{11.131}$$

$$\frac{d\sigma_A}{d\Omega} = \frac{16V_0^2\mu^2a^2}{\hbar^4}\left[\frac{1}{\left(a^2 + 4k^2\sin^2(\theta/2)\right)^2} - \frac{1}{(a^2 + 4k^2\cos^2(\theta/2))^2}\right]^2. \tag{11.132}$$

11.6 Solved Problems

Problem 11.1

(a) Calculate the differential cross section in the Born approximation for the potential $V(r) = V_0e^{-r/R}/r$, known as the Yukawa potential.

(b) Calculate the total cross section.

(c) Find the relation between V_0 and R so that the Born approximation is valid.

Solution

(a) Inserting $V(r) = V_0e^{-r/R}/r$ into (11.70), we obtain

$$\frac{d\sigma}{d\Omega} = \frac{4\mu^2V_0^2}{\hbar^4q^2}\left|\int_0^\infty e^{-r/R}\sin(qr)\,dr\right|^2, \tag{11.133}$$

where

$$\int_0^\infty e^{-r/R}\sin(qr)\,dr = \frac{1}{2i}\int_0^\infty e^{-(1/R-iq)r}\,dr - \frac{1}{2i}\int_0^\infty e^{-(1/R+iq)r}\,dr$$
$$= \frac{1}{2i}\left[\frac{1}{1/R-iq} - \frac{1}{1/R+iq}\right] = \frac{q}{1/R^2+q^2}; \tag{11.134}$$

hence

$$\frac{d\sigma}{d\Omega} = \frac{4\mu^2V_0^2}{\hbar^4}\frac{1}{(1/R^2+q^2)^2} = \frac{4\mu^2V_0^2}{\hbar^4}\frac{1}{\left[1/R^2+4k^2\sin^2(\theta/2)\right]^2}. \tag{11.135}$$

Note that a connection can be established between this relation and the differential cross section for a Coulomb potential $V(r) = Z_1Z_2e^2/r$. For this, we need only to insert $V_0 = -Z_1Z_2e^2$ into (11.135) and then take the limit $R \to \infty$; this leads to (11.77):

$$\lim_{R\to\infty}\left(\frac{d\sigma}{d\Omega}\right)_{Yukawa} = \left(\frac{\mu V_0}{2\hbar^2k^2}\right)^2\frac{1}{\sin^4(\theta/2)}$$
$$= \left(\frac{Z_1Z_2\mu e^2}{2\hbar^2k^2}\right)^2\frac{1}{\sin^4(\theta/2)} = \left(\frac{d\sigma}{d\Omega}\right)_{Rutherford}. \tag{11.136}$$

(b) The total cross section can be obtained at once from (11.135):

$$\sigma = \int\frac{d\sigma}{d\Omega}\sin\theta\,d\theta d\varphi = 2\pi\int_0^\pi\frac{d\sigma}{d\Omega}\sin\theta\,d\theta = 2\pi\frac{4\mu^2V_0^2R^4}{\hbar^4}\int_0^\pi\frac{\sin\theta\,d\theta}{\left(1+4k^2R^2\sin^2(\theta/2)\right)^2}. \tag{11.137}$$

The change of variable $x = 2kR\sin(\theta/2)$ leads to $\sin\theta\,d\theta = x\,dx/(k^2R^2)$; hence

$$\sigma = \frac{8\pi\mu^2V_0^2R^4}{\hbar^4}\frac{1}{k^2R^2}\int_0^{2kR}\frac{x\,dx}{(1+x^2)^2} = \frac{16\pi\mu^2V_0^2R^4}{\hbar^4}\frac{1}{1+4k^2R^2}$$
$$= \frac{16\pi\mu^2V_0^2R^4}{\hbar^4}\frac{1}{1+8\mu ER^2/\hbar^2}, \tag{11.138}$$

where we have used $k^2 = 2\mu E/\hbar^2$; E is the energy of the scattered particle.

(c) The validity condition of the Born approximation is

$$\frac{\mu V_0}{\hbar^2k^2}\left|\int_0^\infty\frac{e^{-ar}}{r}(e^{2ikr}-1)\,dr\right| \ll 1, \tag{11.139}$$

where $a = 1/R$. To evaluate the integral

$$I = \int_0^\infty\frac{e^{-ar}}{r}(e^{2ikr}-1)\,dr \tag{11.140}$$

let us differentiate it with respect to the parameter a:

$$\frac{\partial I}{\partial a} = -\int_0^\infty e^{-ar}(e^{2ikr}-1)\,dr = -\frac{1}{a-2ik}+\frac{1}{a}. \tag{11.141}$$

Now, integrating over the parameter a such that $I(a = +\infty) = 0$, we obtain

$$I = \ln a - \ln(a - 2ik) = -\ln\left(1 - 2i\frac{k}{a}\right) = -\frac{1}{2}\ln(1 + \frac{4k^2}{a^2}) + i\tan^{-1}\left(\frac{2k}{a}\right). \tag{11.142}$$

Thus, the validity condition (11.139) becomes

$$\frac{\mu V_0}{\hbar^2 k^2}\left\{\frac{1}{4}\left[\ln(1 + 4k^2R^2)\right]^2 + \left(\tan^{-1}(2kR)\right)^2\right\}^{1/2} \ll 1. \tag{11.143}$$

Problem 11.2
Find the differential and total cross sections for the scattering of slow (small velocity) particles from a spherical delta potential $V(r) = V_0\delta(r - a)$ (you may use a partial wave analysis). Discuss what happens if there is no scattering potential.

Solution
In the case where the incident particles have small velocities, only the s-waves, $l = 0$, contribute to the scattering. The differential and total cross sections are given for $l = 0$ by (11.104):

$$\frac{d\sigma}{d\Omega} = |f_0|^2 = \frac{1}{k^2}\sin^2\delta_0, \qquad \sigma = 4\pi|f_0|^2 = \frac{4\pi}{k^2}\sin^2\delta_0 \qquad (l = 0). \tag{11.144}$$

We need now to find the phase shift δ_0. For this, we need to consider the Schrödinger equation for the radial function:

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2} + \left[V_0\delta(r - a) + \frac{l(l+1)\hbar^2}{2mr^2}\right]u(r) = Eu(r), \tag{11.145}$$

where $u(r) = rR(r)$. In the case of s states and $r \neq a$, this equation yields

$$\frac{d^2u(r)}{dr^2} = -k^2u(r), \tag{11.146}$$

where $k^2 = 2mE/\hbar^2$. The acceptable solutions of this equation must vanish at $r = 0$ and be finite at $r \to \infty$:

$$u(r) = \begin{cases} u_1(r) = A\sin(kr), & 0 < r < a, \\ u_2(r) = B\sin(kr + \delta_0), & r > a. \end{cases} \tag{11.147}$$

The continuity of $u(r)$ at $r = a$, $u_2(a) = u_1(a)$, leads to

$$B\sin(ka + \delta_0) = A\sin(ka). \tag{11.148}$$

On the other hand, integrating (11.145) (with $l = 0$) from $r = a - \varepsilon$ to $r = a + \varepsilon$, we obtain

$$-\frac{\hbar^2}{2m}\int_{a-\varepsilon}^{a+\varepsilon}\frac{d^2u(r)}{dr^2}\,dr + V_0\int_{a-\varepsilon}^{a+\varepsilon}\delta(r - a)u(r)\,dr = E\int_{a-\varepsilon}^{a+\varepsilon}u(r)\,dr, \tag{11.149}$$

and taking the limit $\varepsilon \to 0$, we end up with

$$\left.\frac{du_2(r)}{dr}\right|_{r=a} - \left.\frac{du_1(r)}{dr}\right|_{r=a} - \frac{2mV_0}{\hbar^2}u_2(a) = 0. \tag{11.150}$$

An insertion of $u_1(r)$ and $u_2(r)$ as given by (11.147) into (11.150) leads to

$$B\left[k\cos(ka+\delta_0)-\frac{2mV_0}{\hbar^2}\sin(ka+\delta_0)\right]=Ak\cos(ka). \tag{11.151}$$

Dividing (11.151) by (11.148), we obtain

$$k\cot(ka+\delta_0)-\frac{2mV_0}{\hbar^2}=k\cot(ka) \quad\Longrightarrow\quad \tan(ka+\delta_0)=\left[\frac{1}{\tan(ka)}+\frac{2mV_0}{k\hbar^2}\right]^{-1}. \tag{11.152}$$

This equation shows that, when there is no scattering potential, $V_0 = 0$, the phase shift is zero, since $\tan(ka+\delta_0) = \tan(ka)$. In this case, equations (11.103) and (11.104) imply that the scattering amplitude and the cross sections all vanish.

If the incident particles have small velocities, $ka \ll 1$, we have $\tan(ka) \simeq ka$ and $\tan(ka+\delta_0) \simeq \tan(\delta_0)$. In this case, equation (11.152) yields

$$\tan\delta_0\simeq\frac{ka}{1+2mV_0a/\hbar^2} \quad\Longrightarrow\quad \sin^2\delta_0\simeq\frac{k^2a^2}{k^2a^2+(1+2mV_0a/\hbar^2)^2}. \tag{11.153}$$

Inserting this relation into (11.144), we obtain

$$\frac{d\sigma}{d\Omega_0}\simeq\frac{a^2}{k^2a^2+(1+2mV_0a/\hbar^2)^2}, \qquad \sigma_0\simeq\frac{4\pi a^2}{k^2a^2+(1+2mV_0a/\hbar^2)^2}. \tag{11.154}$$

Problem 11.3
Consider the scattering of a particle of mass m from a hard sphere potential: $V(r) = \infty$ for $r < a$ and $V(r) = 0$ for $r > a$.

(a) Calculate the total cross section in the low-energy limit. Find a numerical estimate for the cross section for the case of scattering 5 keV protons from a hard sphere of radius $a = 6$ fm.

(b) Calculate the total cross section in the high-energy limit. Find a numerical estimate for the cross section for the case of 700 MeV protons with $a = 6$ fm.

Solution

(a) As the scattering is dominated at low energies by s-waves, $l = 0$, the radial Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2}=Eu(r) \qquad (r>a), \tag{11.155}$$

where $u(r) = rR(r)$. The solutions of this equation are

$$u(r)=\begin{cases} u_1(r)=0, & r<a,\\ u_2(r)=A\sin(kr+\delta_0), & r>a,\end{cases} \tag{11.156}$$

where $k^2 = 2mE/\hbar^2$. The continuity of $u(r)$ at $r = a$ leads to

$$\sin(ka+\delta_0)=0 \quad\Longrightarrow\quad \tan\delta_0=-\tan(ka) \quad\Longrightarrow\quad \sin^2\delta_0=\sin^2(ka), \tag{11.157}$$

since $\sin^2\alpha = 1/(1+\cot^2\alpha)$. The lowest value of the phase shift is $\delta_0 = -ka$; it is negative, as it should be for a repulsive potential. An insertion of $\sin^2\delta_0 = \sin^2(ka)$ into (11.104) yields

$$\sigma_0 = \frac{4\pi}{k^2}\sin^2\delta_0 = \frac{4\pi}{k^2}\sin^2(ka). \tag{11.158}$$

For low energies, $ka \ll 1$, we have $\sin(ka) \simeq ka$ and hence $\sigma_0 \simeq 4\pi a^2$, which is four times the classical value πa^2.

To obtain a numerical estimate of (11.158), we need first to calculate k^2. For this, we need simply to use the relation $E = \hbar^2k^2/(2m_p) = 5\,\text{keV}$, since the proton moves as a free particle before scattering. Using $m_pc^2 = 938.27\,\text{MeV}$ and $\hbar c = 197.33\,\text{MeV fm}$, we have

$$k^2 = \frac{2m_pE}{\hbar^2} = \frac{2(m_pc^2)E}{(\hbar c)^2} = \frac{2(939.57\,\text{MeV})(5\times10^{-3}\,\text{MeV})}{(197.33\,\text{MeV fm})^2} = 0.24\times10^{-3}\,\text{fm}^{-2}. \tag{11.159}$$

Thus $k = 0.0155\,\text{fm}^{-1}$; the wave shift is given by $\delta_0 = -ka = -0.093\,\text{rad} = -5.33°$. Inserting these values into (11.158), we obtain

$$\sigma = \frac{4\pi}{0.24\times10^{-3}\,\text{fm}^{-2}}\sin^2(5.33) = 449.89\,\text{fm}^2 = 4.5\,\text{barn}. \tag{11.160}$$

(b) In the high-energy limit, $ka \gg 1$, the number of partial waves contributing to the scattering is large. Assuming that $l_{max} \simeq ka$, we may rewrite (11.102) as

$$\sigma = \frac{4\pi}{k^2}\sum_{l=0}^{l_{max}}(2l+1)\sin^2\delta_l. \tag{11.161}$$

Since so many values of l contribute in this relation, we may replace $\sin^2\delta_l$ by its average value, $\frac{1}{2}$; hence

$$\sigma \simeq \frac{4\pi}{k^2}\frac{1}{2}\sum_{l=0}^{l_{max}}(2l+1) = \frac{2\pi}{k^2}(l_{max}+1)^2, \tag{11.162}$$

where we have used $\sum_{l=0}^{n}(2l+1) = (n+1)^2$. Since $l_{max} \gg 1$ we have

$$\sigma \simeq \frac{2\pi}{k^2}l_{max}^2 = \frac{2\pi}{k^2}(ka)^2 = 2\pi a^2. \tag{11.163}$$

Since $a = 6\,\text{fm}$, we have $\sigma \simeq 2\pi(6\,\text{fm})^2 = 226.1\,\text{fm}^2 = 2.26\,\text{barn}$. This is almost half the value obtained in (11.160).

In conclusion, the cross section from a hard sphere potential is (a) four times the classical value, πa^2, for low-energy scattering and (b) twice the classical value for high-energy scattering.

Problem 11.4

Calculate the total cross section for the low-energy scattering of a particle of mass m from an attractive square well potential $V(r) = -V_0$ for $r < a$ and $V(r) = 0$ for $r > a$, with $V_0 > 0$.

Solution
Since the scattering is dominated at low energies by the s partial waves, $l = 0$, the Schrödinger equation for the radial function is given by

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2} - V_0 u(r) = Eu(r) \qquad (r < a), \tag{11.164}$$

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2} = Eu(r) \qquad (r > a), \tag{11.165}$$

where $u(r) = rR(r)$. The solutions of these equations for positive energy states are

$$u(r) = \begin{cases} u_1(r) = A\sin(k_1 r), & r < a, \\ u_2(r) = B\sin(k_2 r + \delta_0), & r > a, \end{cases} \tag{11.166}$$

where $k_1^2 = 2m(E + V_0)/\hbar^2$ and $k_2^2 = 2mE/\hbar^2$. The continuity of $u(r)$ and its first derivative, $u'(r) = du(r)/dr$, at $r = a$ yield

$$\left.\frac{u_2(r)}{u_2'(r)}\right|_{r=a} = \left.\frac{u_1(r)}{u_1'(r)}\right|_{r=a} \quad \Longrightarrow \quad \frac{1}{k_2}\tan(k_2 a + \delta_0) = \frac{1}{k_1}\tan(k_1 a), \tag{11.167}$$

which yields

$$\delta_0 = -k_2 a + \tan^{-1}\left[\frac{k_2}{k_1}\tan(k_1 a)\right]. \tag{11.168}$$

Since

$$\tan(k_2 a + \delta_0) = \frac{\sin(k_2 a)\cos\delta_0 + \cos(k_2 a)\sin\delta_0}{\cos(k_2 a)\cos\delta_0 - \sin(k_2 a)\sin\delta_0} = \frac{\tan(k_2 a) + \tan\delta_0}{1 - \tan(k_2 a)\tan\delta_0}, \tag{11.169}$$

we can reduce Eq. (11.167) to

$$\tan\delta_0 = \frac{k_2\tan(k_1 a) - k_1\tan(k_2 a)}{k_1 + k_2\tan(k_1 a)\tan(k_2 a)}. \tag{11.170}$$

Using the relation $\sin^2\delta_0 = 1/(1 + 1/\tan^2\delta_0)$, we can write

$$\sin^2\delta_0 = \left[1 + \left(\frac{k_1 + k_2\tan(k_1 a)\tan(k_2 a)}{k_2\tan(k_1 a) - k_1\tan(k_2 a)}\right)^2\right]^{-1}, \tag{11.171}$$

which, when inserted into (11.104), leads to

$$\sigma_0 = \frac{4\pi}{k_1^2}\sin^2\delta_0 = \frac{4\pi}{k_1^2}\left[1 + \left(\frac{k_1 + k_2\tan(k_1 a)\tan(k_2 a)}{k_2\tan(k_1 a) - k_1\tan(k_2 a)}\right)^2\right]^{-1}. \tag{11.172}$$

If $k_2 a \ll 1$ then (11.170) becomes $\tan\delta_0 \simeq \frac{\tan(k_1 a) - k_1 a}{k_1/k_2 + k_2 a\tan(k_1 a)}$, since $\tan(k_2 a) \simeq k_2 a$. Thus, if $k_2 a \ll 1$ and if E (the scattering energy) is such that $\tan(k_1 a) \simeq k_1 a$, we have $\tan\delta_0 = 0$; hence there will be no s-wave scattering and the cross section vanishes. Note that if the square well potential is extended to a hard sphere potential, i.e., $E \to 0$ and $V_0 \to \infty$, equation (11.168) yields the phase shift of scattering from a hard sphere $\delta_0 = -ka$, since $(k_2/k_1)\tan(k_1 a) \to 0$.

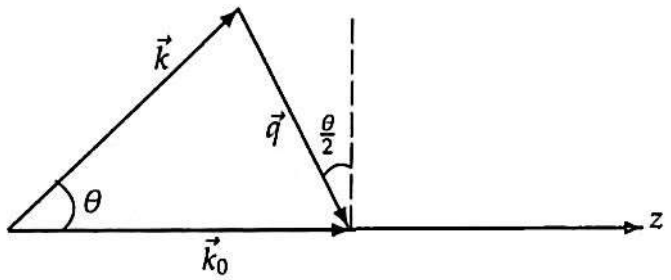


Figure 11.8: Particle traveling initially along the z-axis (taken here horizontally) scatters at an angle θ, with $q = |\vec{k_0} - \vec{k}| = 2k\sin(\theta/2)$, since $k_0 = k$ and $q_z = q\sin(\theta/2)$.

Problem 11.5
Find the differential and total cross sections in the first Born approximation for the elastic scattering of a particle of mass m, which is initially traveling along the z-axis, from a non-spherical, double-delta potential $V(\vec{r}) = V_0\delta(\vec{r} - a\vec{k}) + V_0\delta(\vec{r} + a\vec{k})$, where $\vec{k}$ is the unit vector along the z-axis.

Solution
Since $V(\vec{r})$ is not spherically symmetric, the differential cross section can be obtained from (11.66):

$$\frac{d\sigma}{d\Omega} = \frac{m^2}{4\pi^2\hbar^4}\left|\int V_0\left[\delta(\vec{r} - a\vec{k}) + \delta(\vec{r} + a\vec{k})\right]e^{i\vec{q}\cdot\vec{r}}\,d^3r\right|^2 = \frac{m^2V_0}{4\pi^2\hbar^4}\,|I|^2\,. \tag{11.173}$$

Since $\delta(\vec{r} \pm a\vec{k}) = \delta(x)\delta(y)\delta(z \pm a)$ we can write the integral I as

$$\begin{aligned} I &= \int \delta(x)e^{ixq_x}\,dx \int \delta(y)e^{iyq_y}\,dy \int \left[\delta(z-a) + \delta(z+a)\right]e^{izq_z}\,dz \\ &= e^{iaq_z} + e^{-iaq_z} = 2\cos(aq_z). \end{aligned} \tag{11.174}$$

The calculation of q_z is somewhat different from that shown in (11.67). Since the incident particle is initially traveling along the z-axis, and since it scatters elastically from the potential $V(\vec{r})$, the magnitudes of its momenta before and after collision are equal. So, as shown in Figure 11.8, we have $q_z = q\sin(\theta/2) = 2k\sin^2(\theta/2)$, since $q = |\vec{k_0} - \vec{k}| = 2k\sin(\theta/2)$. Thus, inserting $I = 2\cos(aq_z) = 2\cos\left[2ak\sin^2(\theta/2)\right]$ into (11.173), we obtain

$$\frac{d\sigma}{d\Omega} = \frac{m^2V_0}{\pi^2\hbar^4}\cos^2\left(2ak\sin^2\frac{\theta}{2}\right). \tag{11.175}$$

The total cross section can be obtained at once from (11.175):

$$\begin{aligned} \sigma &= \int \frac{d\sigma}{d\Omega}\sin\theta\,d\theta d\varphi = 2\pi\int_0^\pi \frac{d\sigma}{d\Omega}\sin\theta\,d\theta \\ &= 2\pi\frac{m^2V_0}{\pi^2\hbar^4}\int_0^\pi \sin\theta\cos^2\left(2ak\sin^2\frac{\theta}{2}\right)d\theta, \end{aligned} \tag{11.176}$$

which, when using the change of variable $x = 2ak\sin^2(\theta/2)$ with $dx = 2ak\sin(\theta/2)\cos(\theta/2)\,d\theta$, leads to

$$\begin{aligned}\sigma &= \frac{2m^2V_0}{\pi\hbar^4}\int_0^{\pi} 2\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)\cos^2\left(2ak\sin^2\frac{\theta}{2}\right)d\theta \\ &= \frac{2m^2V_0}{\pi ak\hbar^4}\int_0^{1}\cos^2(x)\,dx \\ &= \frac{m^2V_0}{\pi ak\hbar^4}\int_0^{1}[1+\cos(2x)]\,dx \\ &= \frac{m^2V_0}{\pi ak\hbar^4}. \end{aligned} \tag{11.177}$$

Problem 11.6

Consider the elastic scattering of 50 MeV neutrons from a nucleus. The phase shifts measured in this experiment are $\delta_0 = 95°$, $\delta_1 = 72°$, $\delta_2 = 60°$, $\delta_3 = 35°$, $\delta_4 = 18°$, $\delta_5 = 5°$; all other phase shifts are negligible (i.e., $\delta_l \simeq 0$ for $l \geq 6$).

(a) Find the total cross section.

(b) Estimate the radius of the nucleus.

Solution

(a) As $\delta_l \simeq 0$ for $l \geq 6$, equation (11.102) yields

$$\begin{aligned}\sigma &= \frac{4\pi}{k^2}\sum_{l=0}^{6}(2l+1)\sin^2\delta_l \\ &= \frac{4\pi}{k^2}\left(\sin^2\delta_0 + 3\sin^2\delta_1 + 5\sin^2\delta_2 + 7\sin^2\delta_3 + 9\sin^2\delta_4 + 11\sin^2\delta_5\right) = \frac{4\pi}{k^2}\times 10.702. \end{aligned} \tag{11.178}$$

To calculate k^2, we need simply to use the relation $E = \hbar^2k^2/(2m_n) = 50\,\text{MeV}$, since the neutrons move as free particles before scattering. Using $m_nc^2 = 939.57\,\text{MeV}$ and $\hbar c = 197.33\,\text{MeV fm}$, we have

$$k^2 = \frac{2m_nE}{\hbar^2} = \frac{2(m_nc^2)E}{(\hbar c)^2} = \frac{2(939.57\,\text{MeV})(50\,\text{MeV})}{(197.33\,\text{MeV fm})^2} = 2.41\,\text{fm}^{-2}. \tag{11.179}$$

An insertion of (11.179) into (11.178) leads to

$$\sigma = \frac{4\pi}{2.41\,\text{fm}^{-2}}\times 10.702 = 55.78\,\text{fm}^2 = 0.558\,\text{barn}. \tag{11.180}$$

(b) At large values of l, when the neutron is at its closest approach to the nucleus, it feels mainly the effect of the centrifugal potential $l(l+1)\hbar^2/(2m_nr^2)$; the effect of the nuclear potential is negligible. We may thus use the approximations $E \simeq l(l+1)\hbar^2/(2m_nr_c^2) \simeq 42\hbar^2/(2m_nr_c^2)$ where we have taken $l \simeq 6$, since $\delta_l \simeq 0$ for $l \geq 6$. A crude value of the radius of the nucleus is then given by

$$r_c \simeq \sqrt{\frac{21\hbar^2}{m_nE}} = \sqrt{\frac{21(\hbar c)^2}{(m_nc^2)E}} = \sqrt{\frac{21(197.33\,\text{MeV fm})^2}{(939.57\,\text{MeV})(50\,\text{MeV})}} = 4.17\,\text{fm}. \tag{11.181}$$

Problem 11.7
Consider the elastic scattering of an electron from a hydrogen atom in its ground state. If the atom is assumed to remain in its ground state after scattering, calculate the differential cross section in the case where the effects resulting from the identical nature of the electrons (a) are ignored and (b) are taken into account (in part (b), discuss the three cases when the electrons are in (i) a spin singlet state, (ii) a spin triplet state, or (iii) an unpolarized state).

Solution
(a) By analogy with (11.63) we may write the differential cross section for this process as

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 = \left|-\frac{\mu}{2\pi\hbar^2}\langle\Psi_f \mid \hat{V} \mid \Psi_i\rangle\right|^2, \tag{11.182}$$

where $\mu \simeq m_e/2$, since this problem can be viewed as the scattering of a particle whose reduced mass is half that of the electron. Assuming the atom to be very massive and that it remains in its ground state after scattering, the initial and final states of the system (incident electron plus the atom) are given by $\Psi_i(\vec{r}, \vec{k}_0, \vec{r}') = e^{i\vec{k}_0\cdot\vec{r}}\psi_0(\vec{r}')$ and $\Psi_f(\vec{r}, \vec{k}, \vec{r}') = e^{i\vec{k}\cdot\vec{r}}\psi_0(\vec{r}')$, where $e^{i\vec{k}_0\cdot\vec{r}}$ and $e^{i\vec{k}\cdot\vec{r}}$ are the states of the incident electron before and after scattering, and $\psi_0(\vec{r}') = (\pi a_0^3)^{-1/2}e^{-r'/a_0}$ is the atom's wave function. We have assumed here that the nucleus is located at the origin and that the position vectors of the incident electron and the atom's electron are given by $\vec{r}$ and $\vec{r}'$, respectively. Since the incident electron experiences an attractive Coulomb interaction $-e^2/r$ with the nucleus and a repulsive interaction $e^2/|\vec{r}-\vec{r}'|$ with the hydrogen's electron, we have

$$f(\theta) = -\frac{\mu}{2\pi\hbar^2}\int d^3r e^{i\vec{q}\cdot\vec{r}}\int d^3r'\,\psi_0^*(\vec{r}')\left[-\frac{e^2}{r} + \frac{e^2}{|\vec{r}-\vec{r}'|}\right]\psi_0(\vec{r}'), \tag{11.183}$$

with $q = |\vec{k}_0 - \vec{k}| = 2k\sin(\theta/2)$, since $k = k_0$ (elastic scattering). Using $\int_0^\infty \sin(qr)\,dr = 1/q$ (see (11.76)), and since $\int_0^\pi e^{iqr\cos\theta}\sin\theta\,d\theta = \int_{-1}^1 e^{iqrx}dx = (2/qr)\sin(qr)$, we obtain the following relation:

$$\int d^3r\frac{e^{i\vec{q}\cdot\vec{r}}}{r} = \int_0^\infty r\,dr\int_0^\pi e^{iqr\cos\theta}\sin\theta\,d\theta\int_0^{2\pi}d\varphi = \frac{4\pi}{q}\int_0^\infty dr\sin(qr) = \frac{4\pi}{q^2}, \tag{11.184}$$

which, when inserted into (11.183) and since $\int d^3r'\,\psi_0^*(\vec{r}')\psi_0(\vec{r}') = 1$, leads to

$$f(\theta) = \frac{\mu}{2\pi\hbar^2}\left[\frac{4\pi e^2}{q^2} - \int d^3r e^{i\vec{q}\cdot\vec{r}}\int d^3r'\,\psi_0^*(\vec{r}')\frac{e^2}{|\vec{r}-\vec{r}'|}\psi_0(\vec{r}')\right]. \tag{11.185}$$

By analogy with (11.184), we have $\int d^3r e^{i\vec{q}\cdot|\vec{r}-\vec{r}'|}/|\vec{r}-\vec{r}'| = 4\pi/q^2$; hence we can reduce the integral in (11.185) to

$$\begin{aligned}\int d^3r\,e^{i\vec{q}\cdot\vec{r}}\int d^3r'\,\psi_0^*(\vec{r}')\frac{e^2}{|\vec{r}-\vec{r}'|}\psi_0(\vec{r}') &= e^2\int d^3r'\,\psi_0^*(\vec{r}')e^{i\vec{q}\cdot\vec{r}'}\psi_0(\vec{r}')\int d^3r\frac{e^{i\vec{q}\cdot|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}\\ &= \frac{4\pi e^2}{q^2}\int d^3r'\,\psi_0^*(\vec{r}')e^{i\vec{q}\cdot\vec{r}'}\psi_0(\vec{r}').\end{aligned} \tag{11.186}$$

The remaining integral of (11.186) can, in turn, be written as

$$\int d^3r'\,\psi_0^*(\vec{r}')e^{i\vec{q}\cdot\vec{r}'}\psi_0(\vec{r}') = \frac{1}{\pi a_0^3}\int_0^\infty r'^2 e^{-2r'/a_0}\,dr' \int_0^\pi e^{iqr'\cos\theta'}\sin\theta' d\theta' \int_0^{2\pi} d\varphi'$$

$$= \frac{4}{qa_0^3}\int_0^\infty r' e^{-2r'/a_0}\sin(qr')\,dr' = \left(1+\frac{a_0^2q^2}{4}\right)^{-2}, \tag{11.187}$$

where we have used the expression for $\int_0^\infty re^{-ar}\sin(qr)\,dr$ calculated in (11.130). Inserting (11.187) into (11.186), and the resulting expression into (11.185), we obtain

$$f(\theta) = \frac{2\mu e^2}{\hbar^2 q^2}\left[1-\left(1+\frac{a_0^2q^2}{4}\right)^{-2}\right] = \frac{\mu e^2}{2k^2\hbar^2\sin^2(\theta/2)}\left[1-\left(1+a_0^2k^2\sin^2\frac{\theta}{2}\right)^{-2}\right]. \tag{11.188}$$

We can thus reduce (11.182) to

$$\frac{d\sigma}{d\Omega} = \frac{4\mu^2e^4}{\hbar^4q^4}\left[1-\left(1+\frac{a_0^2q^2}{4}\right)^{-2}\right]^2 = \frac{\mu^2e^4}{4k^4\hbar^4\sin^4\frac{\theta}{2}}\left[1-\left(1+a_0^2k^2\sin^2(\theta/2)\right)^{-2}\right]^2, \tag{11.189}$$

with $q = 2k\sin(\theta/2)$.

(b) (i) If the electrons are in their spin singlet state (antisymmetric), the spatial wave function must be symmetric; hence the differential cross section is

$$\frac{d\sigma_S}{d\Omega} = \left|f(\theta)+f(\pi-\theta)\right|^2, \tag{11.190}$$

where $f(\theta)$ is given by (11.188) and

$$f(\pi-\theta) = \frac{2\mu e^2}{\hbar^2q^2}\left[1-\left(1+\frac{a_0^2q^2}{4}\right)^{-2}\right] = \frac{\mu e^2}{2k^2\hbar^2\cos^2\frac{\theta}{2}}\left[1-\left(1+a_0^2k^2\cos^2(\theta/2)\right)^{-2}\right], \tag{11.191}$$

since $\sin(\pi-\theta/2) = \cos(\theta/2)$.

(ii) If, however, the electrons are in their spin triplet state, the spatial wave function must be antisymmetric; hence

$$\frac{d\sigma_A}{d\Omega} = \left|f(\theta)-f(\pi-\theta)\right|^2. \tag{11.192}$$

(iii) Finally, if the electrons are unpolarized, the differential cross section must be a mixture of (11.191) and (11.192):

$$\frac{d\sigma}{d\Omega} = \frac{1}{4}\frac{d\sigma_S}{d\Omega} + \frac{3}{4}\frac{d\sigma_A}{d\Omega} = \frac{1}{4}\left|f(\theta)+f(\pi-\theta)\right|^2 + \frac{3}{4}\left|f(\theta)-f(\pi-\theta)\right|^2. \tag{11.193}$$

Problem 11.8

In an experiment, 650 MeV π^0 pions are scattered from a heavy, totally absorbing nucleus of radius 1.4 fm.

(a) Estimate the total elastic and total inelastic cross sections.

(b) Calculate the scattering amplitude and check the validity of the optical theorem.

(c) Using the scattering amplitude found in (b), calculate and plot the differential cross section for elastic scattering. Calculate the total elastic cross section and verify that it agrees with the expression found in (a).

Solution

(a) In the case of a totally absorbing nucleus, $\eta_l(k) = 0$, the total elastic and inelastic cross sections, which are given by (11.113) and (11.114), become equal:

$$\sigma_{el} = \frac{\pi}{k^2}\sum_{l=0}^{l_{max}}(2l+1) = \sigma_{inel}. \tag{11.194}$$

This experiment can be viewed as a scattering of high-energy pions, $E = 650\,\text{MeV}$, from a black "disk" of radius $a = 1.4\,\text{fm}$; thus, the number of partial waves involved in this scattering can be obtained from $l_{max} \simeq ka$, where $k = \sqrt{2m_{\pi^0}E/\hbar^2}$. Since the rest mass energy of a π^0 pion is $m_{\pi^0}c^2 \simeq 135\,\text{MeV}$ and since $\hbar c = 197.33\,\text{MeV fm}$, we have

$$k \simeq \sqrt{\frac{2m_{\pi^0}E}{\hbar^2}} = \sqrt{\frac{2(m_{\pi^0}c^2)E}{(\hbar c)^2}} = \sqrt{\frac{2(135\,\text{MeV})(650\,\text{MeV})}{(197.33\,\text{MeV fm})^2}} = 2.12\,\text{fm}^{-1}; \tag{11.195}$$

hence $l_{max} = ka \simeq (2.12\,\text{fm}^{-1})(1.4\,\text{fm}) = 2.97 \simeq 3$. We can thus reduce (11.194) to

$$\sigma_{el} = \sigma_{inel} = \frac{\pi}{k^2}\sum_{l=0}^{3}(2l+1) = \frac{16\pi}{k^2} \simeq \frac{16\pi}{(2.12\,\text{fm}^{-1})^2} = 40.1\,\text{fm}^2 = 0.40\,\text{barn}. \tag{11.196}$$

The total cross section

$$\sigma_{tot} = \sigma_{el} + \sigma_{inel} = \frac{32\pi}{k^2} = 0.80\,\text{barn}. \tag{11.197}$$

(b) The scattering amplitude can be obtained from (11.112) with $\eta_l(k) = 0$:

$$\begin{aligned} f(\theta) &= \frac{i}{2k}\sum_{l=0}^{3}(2l+1)P_l(\cos\theta) \\ &= \frac{i}{2k}\left[1 + 3\cos\theta + \frac{5}{2}(3\cos^2\theta - 1) + \frac{7}{2}(5\cos^3\theta - 3\cos\theta)\right], \end{aligned} \tag{11.198}$$

where we have used the following Legendre polynomials: $P_0(u) = 1$, $P_1(u) = u$, $P_2(u) = \frac{1}{2}(3u^2 - 1)$, $P_3(u) = \frac{1}{2}(5u^3 - 3u)$. The forward scattering amplitude ($\theta = 0$) is

$$f(0) = \frac{i}{2k}\left[1 + 3 + \frac{5}{2}(3-1) + \frac{7}{2}(5-3)\right] = \frac{8i}{k}. \tag{11.199}$$

Combining (11.197) and (11.199), we get the optical theorem: $\text{Im}f(0) = (k/4\pi)\sigma_{tot} = 8/k$.

(c) From (11.198) the differential elastic cross section is

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 = \frac{1}{4k^2}\left[1 + 3\cos\theta + \frac{5}{2}(3\cos^2\theta - 1) + \frac{7}{2}(5\cos^3\theta - 3\cos\theta)\right]^2. \tag{11.200}$$

As shown in Figure 11.9, the differential cross section displays an interference pattern due to the superposition of incoming and outgoing waves. The total elastic cross section is given by $\sigma_{el} = \int_0^\pi |f(\theta)|^2 \sin\theta\, d\theta \int_0^{2\pi} d\varphi$ which, combined with (11.200), leads to

$$\sigma_{el} = \frac{2\pi}{4k^2}\int_0^\pi \left[1 + 3\cos\theta + \frac{5}{2}(3\cos^2\theta - 1) + \frac{7}{2}(5\cos^3\theta - 3\cos\theta)\right]^2 \sin\theta\, d\theta = \frac{16\pi}{k^2}. \tag{11.201}$$

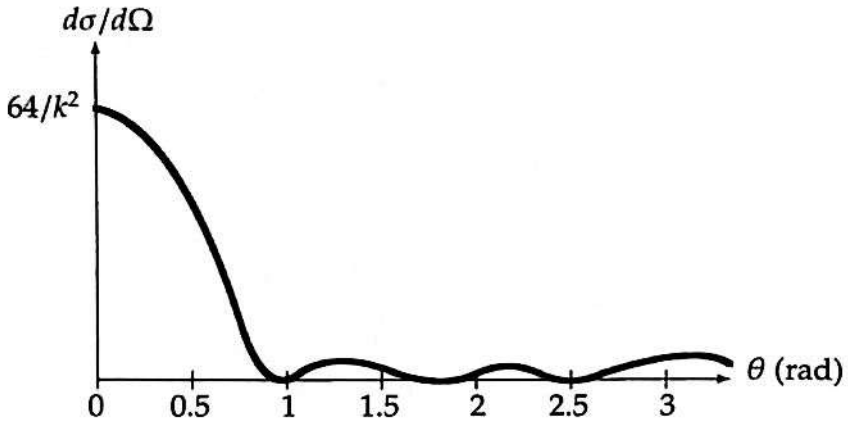


Figure 11.9: Plot of $\frac{d\sigma}{d\Omega} = \frac{1}{4k^2}\left[1 + 3\cos\theta + \frac{5}{2}(3\cos^2\theta - 1) + \frac{7}{2}(5\cos^3\theta - 3\cos\theta)\right]^2$.

This is the same expression we obtained in (11.196). Unlike the differential cross section, the total cross section displays no interference pattern because its final expression does not depend on any angle, since the angles were integrated over.

11.7 Exercises

Exercise 11.1
Consider the scattering of a 5 MeV alpha particle (i.e., a helium nucleus with $Z_1 = 2$ and $A_1 = 4$) from an aluminum nucleus ($Z_2 = 13$ and $A_2 = 27$). If the scattering angle of the alpha particle in the Lab frame is $\theta_1 = 30°$,

(a) find its scattering angle θ in the CM frame and

(b) give a numerical estimate of the Rutherford cross section.

Exercise 11.2

(a) Find the differential and total cross sections for the classical collision of two hard spheres of radius r and R, where R is the radius of the larger sphere; the larger sphere is considered to be stationary.

(b) From the results of (a) find the differential and total cross sections for the scattering of pointlike particles from a hard stationary sphere of radius R. *Hint:* You may use the classical relation $d\sigma/d\Omega = -[b(\theta)/\sin\theta]db/d\theta$, where $b(\theta)$ is the impact parameter.

Exercise 11.3
Consider the scattering from the potential $V(r) = V_0 e^{-r^2/a^2}$. Find

(a) the differential cross section in the first Born approximation and

(b) the total cross section.

Exercise 11.4

Consider the scattering from the δ–function potential $V(r) = V_0\delta(\vec{r})$. Use the first Born approximation to find

(a) the scattering amplitude,

(b) the differential cross section, and

(b) the total cross section.

Exercise 11.5

Calculate the differential cross section in the first Born approximation for the scattering of a particle by an attractive square well potential: $V(r) = -V_0$ for $r < a$ and $V(r) = 0$ for $r > a$, with $V_0 > 0$.

Exercise 11.6

Consider the elastic scattering from the delta potential $V(r) = V_0\delta(r - a)$.

(a) Calculate the differential cross section in the first Born approximation.

(b) Find an expression between V_0, a, μ, and k so the Born approximation is valid.

Exercise 11.7

Consider the elastic scattering from the potential $V(r) = V_0e^{-r/a}$, where V_0 and a are constant.

(a) Calculate the differential cross section in the first Born approximation.

(b) Find an expression between V_0, a, μ, and k so the Born approximation is valid.

(c) Find the total cross section using the Born approximation.

Exercise 11.8

Find the differential cross section in the first Born approximation for the elastic scattering of a particle of mass m, which is initially traveling along the z-axis, from a nonspherical, double-delta potential:

$$V(\vec{r}) = V_0\delta(\vec{r} - a\vec{k}) - V_0\delta(\vec{r} + a\vec{k}),$$

where $\vec{k}$ is the unit vector along the z-axis.

Exercise 11.9

Find the differential cross section in the first Born approximation for neutron–neutron scattering in the case where the potential is approximated by $V(r) = V_0e^{-r/a}$.

Exercise 11.10

Consider the elastic scattering of a particle of mass m and initial momentum $\hbar k$ off a delta potential $V(\vec{r}) = V_0\delta(x)\delta(y)\delta(z - a)$, where V_0 is a constant.

(a) What is the physical dimensions of the constant V_0?

(b) Calculate the differential cross sections in the first Born approximation.

(c) Repeat (b) for the case where the potential is now given by

$$V(\vec{r}) = V_0\delta(x)\left[\delta(y - b)\delta(z) + \delta(y)\delta(z - a)\right].$$

Exercise 11.11

Consider the S-wave ($l = 0$) scattering of a particle of mass m from a repulsive spherical potential $V(r) = V_0$ for $r < a$ and $V(r) = 0$ for $r > a$, with $V_0 > 0$.

(a) Calculate S–wave ($l = 0$) phase shift and the total cross section.

(b) Show that in the limit $V_0 \to \infty$, the phase shift is given by $\delta_0 = -ka$. Find the total cross section.

Exercise 11.12
Consider the S-wave neutron–neutron scattering where the interaction potential is approximated by $V(r) = V_0 \vec{S}_1 \cdot \vec{S}_2 e^{-r/a}$, where $\vec{S}_1$ and $\vec{S}_2$ are the spin vector operators of the two neutrons, and $V_0 > 0$. Find the differential cross section in the first Born approximation.

Exercise 11.13
Consider the S-partial wave scattering ($l = 0$) between two identical spin 1/2 particles where the interaction potential is given approximately by

$$\hat{V}(r) = V_0 \vec{S}_1 \cdot \vec{S}_2 \delta(r - a),$$

where $\vec{S}_1$ and $\vec{S}_2$ are the spin vector operators of the two particles, and $V_0 > 0$. Assuming that the incident and target particles are unpolarized, find the differential and total cross sections.

Exercise 11.14
Consider the elastic scattering of 170 MeV neutrons from a nucleus of radius $a = 1.05$ fm. Consider the hypothetical case where the phase shifts measured in this experiment are given by $\delta_l = \frac{180^\circ}{l+2}$.

(a) Estimate the maximum angular momentum l_{max}.

(b) Find the total cross section.

Chapter 12

Relativistic Quantum Mechanics

12.1 Introduction

Up to this point in the book, we have dealt with nonrelativistic quantum mechanics only. This description, which is founded on the Schrödinger equation, is incomplete. Based on the classical energy–momentum expression $E = P^2/(2m) + V$ for a particle moving in a potential V, nonrelativistic quantum mechanics cannot describe the dynamics of particles moving at speeds close to the speed of light. In addition to this limitation, the Schrödinger[1] equation, originally proposed in 1926, fails to explain many quantum concepts and phenomena such as the electron's spin (it was Pauli[2] who, in 1927, has added it to the Schrödinger equation in a rather ad hoc manner), the fine structure of the hydrogen atom, and, in particular, those phenomena that appear at relativistic speeds such as particle creation, particle annihilation, and particle decays. Apart from the inability of the Schrödinger equation to describe relativistic phenomena, the need to blend relativity and quantum mechanics into a unified theory is quite compelling at least from an aesthetically theoretical perspective.

When does a relativistic quantum mechanical treatment becomes necessary? When the speed of a particle approaches that of light, $p \sim mc$, or when its energy is large compared to its rest-mass energy[3], mc^2, a relativistic treatment mechanics becomes a must. For instance, photons can only be relativistic, since they always travel at the speed of light (due to their zero-rest-mass energy). In contrast, nonrelativistic quantum mechanics deals only with particles having energies and momenta that are much smaller than mc^2 and mc, respectively. So the domain of validity of nonrelativistic quantum mechanics is defined by:

$$p \ll mc, \qquad E \ll mc^2. \tag{12.1}$$

At these low energies, the phenomenon of particle creation cannot occur. Particle creation is a purely relativistic phenomenon; it occurs only at energies much higher than mc^2. In this context, one might mention the 2012 historical experiment that led to the creation and detection of the long-awaited Higgs boson particle, an effect that occurs only at relativistic energies: in the Large Hadron Collider (LHC), the world's largest and most powerful particle

[1] E. Schrödinger, *An Undulatory Theory of the Mechanics of Atoms and Molecules*, Phys. Rev. 28, 1049 (1926).
[2] Wolfgang Pauli *Zur Quantenmechanik des magnetischen Elektrons Zeitschrift für Physik* 43, 601–623 (1927).
[3] When a particle is at rest, its energy is expressed as mc^2 which is known as its rest-mass energy.

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

collider, two beams of protons are collided after being accelerated to energies of about 7 TeV (1 TeV = 10^{12} eV = 10^3 GeV) per proton (7 TeV is about 7.5×10^3 larger than the proton's rest-mass energy of 0.94 GeV); the speed of protons at 7 TeV is about $0.999999991c$. These high energy collisions led to the creation of the Higgs boson particle, with a rest-mass energy of about 125 GeV. At these high energies, nonrelativistic quantum mechanics is simply helpless in providing a description of the phenomena observed in the LHC experiment.

It should be noted that we can use the Heisenberg uncertainty principle to offer another perspective to determine the conditions that necessitate a relativistic description. From conditions (12.1) and the uncertainty principle, $\Delta x \Delta p \geq \hbar$, we can write

$$\Delta x \Delta p \geq \hbar \qquad \Longrightarrow \qquad \Delta x \geq \frac{\hbar}{\Delta p} \gg \lambda_c, \tag{12.2}$$

where $\lambda_c = \hbar/(mc)$ is the Compton wavelength of the particle[4]. This implies that the position of a relativistic particle cannot be determined more accurately than a distance which is large compared to its Compton wavelength. So, the Compton wavelength is the length scale at which we need to utilize a relativistic description; at this scale, the processes of particle creation become possible. This can be viewed as follows: if we use a radiation of wavelength λ to localize a particle of mass m and if $\lambda < \lambda_c$, the radiation's photons have more energy than mc^2 (the rest-mass energy of the particle). In this case, the radiative photons will most likely lead to the creation of more particles of mass m; this will make it almost impossible to localize the particle.

From a historical perspective, a number of attempts were undertaken between 1926–1930 to construct a quantum theory that accounts for relativistic effects so as to address the limitations of the pre-1926 nonrelativistic quantum mechanics; these efforts have focused primarily on finding a relativistic generalization of the Schrödinger equation. In our presentation, we follow a roughly historical order so as not to miss the difficulties that were initially encountered and the ideas that were developed to overcome them. These attempts have broadly led to two relativistic equations:

- An equation that was introduced independently by Schrödinger (1926), Gordon[5] (1926), and by Klein[6] (1927), but later became known as the Klein–Gordon equation and has since been found to describe spin-zero particles such as π–mesons. As we will see in the Historical Note of Section 12.3, Schrödinger was the first to have derived this equation but was dissatisfied with it and never published it due to its various untractable problems such as negative energies and negative probability densities.

- An equation due to Dirac[7] (1928) that describes spin-1/2 particles such as electrons.

As will be explained below, while the Klein–Gordon equation quickly ran into some untractable conceptual problems (see Subsection 12.3.4), the Dirac equation turned out to be very successful on two major fronts: (a) it managed to avoid the problems that have plagued the Klein–Gordon equation and (b) it succeeded in addressing a number of limitations of the

[4]The Compton wavelength of a particle is defined as the wavelength of a photon whose energy, $h\nu \equiv hc/\lambda$, is equal to the rest energy of the particle, mc^2; that is: $mc^2 = h\nu \equiv hc/\lambda \implies \lambda = \hbar/(mc)$.

[5]W. Gordon, *Z. Physik*, **40**, 117 (1926).

[6]O. Klein, *Z. Physik*, **37**, 895 (1926); **41**, 407 (1927).

[7]P.A.M. Dirac, *Proc. Roy. Soc.* (London) **A117**, 610 (1928); *ibid.* **A118**, 351 (1928).

Schrödinger equation, most notably the proper description of the motion of an electron, an accurate treatment of the fine structure of the hydrogen atom[8], the prediction of the existence of antiparticles, and the electron's spin. While the Schrödinger equation fails to deal with the electron's spin from first principles, it emerges naturally from Dirac's equation; the electron's spin is built into the Dirac equation from the very outset. Due to its unassailable success, the Dirac equation became the foundation of relativistic quantum mechanics as well as quantum field theory.

In this chapter, we cover some of the most salient elements of relativistic quantum mechanics. We begin our exposition with a brief outline of the relativistic notation which will render the formalism transparent and easy to follow; an essential property of any relativistic quantum theory is the need to be consistent with the Lorentz transformations of special relativity. After discussing the Klein–Gordon equation and its internal inconsistencies, we present the most essential aspects of the Dirac Equation as well as some of its successes in accounting for a number of relativistic quantum phenomena. We conclude this chapter by offering a few comments on what lies beyond the Klein–Gordon and Dirac equations and how their limitations are addressed adequately by the ultimate quantum theory of relativistic phenomena: Quantum Field Theory.

12.2 Four Vectors and Minkowski Metric

In this section, we present a brief outline of those tools that will make it easy to follow the formalisms of the Klein–Gordon and Dirac equations. A somewhat more expanded treatment of these tools is available in appendices D and E.

12.2.1 Euclidean Metric

As discussed in Appendix D, where we used the *index notation*, we can express any vector $\vec{A}$ in the 3-dimentional Euclidean space in terms of the three basis vectors $(\hat{e}_x, \hat{e}_x, \hat{e}_z) \equiv (\hat{e}_1, \hat{e}_2, \hat{e}_3)$ as follows

$$\vec{A} = A^1\hat{e}_1 + A^2\hat{e}_2 + A^3\hat{e}_3 = \sum_{i=1}^{3} A^i\hat{e}_i \equiv A^i\hat{e}_i, \tag{12.3}$$

where we have used the *Einstein summation convention* in the last step. According to this convention, whenever an index variable (here "i") appears twice in a single term, summation of that term over all the values of the index is automatically implied (here $i = 1, 2, 3$), even though the summation symbol is not explicitly written. This is one of the advantages for using the index notation. For notational brevity, we will be using the Einstein summation convention throughout Chapter 12 and the relevant appendices, D, E, F, and G.

Distance, or equivalently the length of a vector, is an invariant irrespective of the reference in which we measure it. For instance, as we saw in Chapter 7 (see Eq. (7.10) on Page 436), the

[8]The energy levels of a hydrogen atom calculated from the nonrelativistic Schrödinger equation miss the contributions from the fine structure splitting; this splitting is due to the relativistic variation of mass and to the electron's spin. The variation of the mass can be explained by the Klein–Gordon equation for spin-zero particles but would not give the correct results for the phenomena which depend on the spin of the electron such as the fine structure and the Zeeman effect. Dirac equation, however, is successful in dealing with the electron's spin and the phenomena involving the spin.

square of a vector $\vec{A}$ is equal to its rotated counterpart $\vec{A}'$

$$\vec{A'}^2 = \vec{A}^2 \quad \longrightarrow \quad (A'^1)^2 + (A'^2)^2 + (A'^3)^2 = (A^1)^2 + (A^2)^2 + (A^3)^2. \tag{12.4}$$

Similarly, the scalar product between any two vectors is invariant under rotations of the coordinate system: :

$$\vec{A'} \cdot \vec{B'} = \vec{A} \cdot \vec{B} \quad \longrightarrow \quad A'^1 B'^1 + A'^2 B'^2 + A'^3 B'^3 = A^1 B^1 + A^2 B^2 + A^3 B^3. \tag{12.5}$$

A useful geometrical tool to deal with distance in a given space is called the *metric* tensor g that operates as follows: when g is "fed" with two vectors, it gives us back their scalar product:

$$g(\vec{A}, \vec{A}) = \vec{A} \cdot \vec{A} = \vec{A}^2 = (A^1)^2 + (A^2)^2 + (A^3)^2, \qquad g(\vec{A}, \vec{B}) = \vec{A} \cdot \vec{B} = A^1 B^1 + A^2 B^2 + A^3 B^3. \tag{12.6}$$

In particular, applying the metric to the Euclidean basis vectors $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$, we obtain

$$g(\hat{e}_i, \hat{e}_j) = \hat{e}_i \cdot \hat{e}_j = g_{ij} = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}, \qquad i, j = 1,2,3, , \tag{12.7}$$

where δ_{ij} is the 3-dimensional *Kronecker delta* and where we have used the obvious fact that the basis vectors $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$ are orthonormal. From this expression, we can easily infer the matrix of the *Euclidean metric*, which turns out to be equal to the unity matrix:

$$[g] = (g_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv [\delta] = (\delta_{ij}). \tag{12.8}$$

Note: In this text, we use $[g]$ or (g_{ij}) to denote the matrix of the metric, while g_{ij} represents the ij element of the metric; the same notation for the Kronecker delta function where $[\delta]$ or (δ_{ij}) represent the matrix and δ_{ij} the ij element of the $\delta-$matrix.

12.2.2 Minkowski Spacetime and the Lorentz Transformation

In classical physics, we use the standard 3-dimensional Euclidean space to describe the motion of objects; in this space, time and the spatial degrees of freedom are totally independent. In special relativity, however, we use a different space where the time and spatical degrees of freedom are interdependent; this is a space in which the rules of special relativity hold. Minkowski has shown in 1908 that the theory of relativity can be best described within the framework of a *four dimensional space*, known ever since then as the *Minkowski spacetime*[9]. Unlike the Euclidean space, which deals with the spatial coordinates only, the Minkowski spacetime includes time in addition to the spatial coordinates in a way that the 3 spatial coordinates and the 1 time component are not separate entities, they are inseparably connected.

A point in the Minknowski spacetime is characterized by what happens or takes place there; this is called an *"event"*. For example, two particles colliding at a certain point in space is an event; this event takes place at a specific place (in the 3-dimensional space) and a unique time seen by an observer in a given reference frame. As a result, an event in

[9] Hermann Minkowski, *Space and Time*, *Physikalische Zeitschrift*, **10**, pp. 75–88 (1908).

Minkowski spacetime is defined or indexed by *four components: three spatial coordinates and one time component.* As a result, the coordinates of an event in the Minkowski spacetime is denoted by the 4-vector $(ct, \vec{r}) \equiv (ct, x, y, z)$ where we have multiplied the time with the speed of light c in the first component (i.e., ct) to ensure that all such four components have the same physical dimensions of length.

One of the most important features of 4-vectors is the fact that they lend themselves quite easily to Lorentz transformations which regulate how their components transform between inertial frames.

Lorentz Transformation

The *Lorentz transformation* deals with one central issue: How to connect the coordinates of an event that are recorded by one observer to the coordinates of the same event recorded by another observer. As shown in Appendix E, the coordinates (ct, x, y, z) and (ct', x', y', z') of an event in two *inertial frames*[10] F and F', where F' is in uniform motion along the x-direction with constant velocity $\vec{v}$ relative to F (see Figure E.1), are related by the *Lorentz Transformation* as follows:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = [\Lambda] \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma(ct - \beta x) \\ \gamma(x - c\beta t) \\ y \\ z \end{pmatrix} \quad \longrightarrow \quad [\Lambda] = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{12.9}$$

where $\beta = v/c$, $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ is the *Lorentz factor*, and $[\Lambda]$ is the matrix of the Lorentz transformation. This is also known as a *boost transformation* along the x-axis.

As shown in Problem 12.8, the Lorentz transformation has an important property: it conserves the length of 4-vectors:

$$\boxed{(ct')^2 - (x'^2 + y'^2 + z'^2) = (ct)^2 - (x^2 + y^2 + z^2).} \tag{12.10}$$

12.2.3 Four-Vectors

The use of 4-vectors can be greatly simplified by adopting the following notations[11]: $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$; that is, we can write an event's 4-displacement vector in the following simplified form that unifies the time and space components:

$$x^\mu = \left(x^0, x^1, x^2, x^3\right) = (ct, x, y, z) \equiv (ct, \vec{r}), \qquad \mu = 0, 1, 2, 3. \tag{12.11}$$

By analogy, we may mention other familiar examples of 4-vectors such as the 4-momentum $p^\mu = (E/c, \vec{p})$ where energy and momentum are also treated on equal footing, the 4-current density $j^\mu = (c\rho, \vec{j})$ involving the charge density $\rho(\vec{r}, t)$ and the current density $\vec{j}(\vec{r}, t)$, the 4-potential $A^\mu = (A^0, \vec{A}) = \left(U/c, \vec{A}\right)$ where U is the electric potential and $\vec{A}$ is the vector potential, the 4-wave vector $k^\mu = p^\mu/\hbar = (E/(\hbar c), \vec{p}/\hbar)$, and so on. For completeness, Table 12.1 offers a

[10] Frames that have no acceleration relative to one another are known as inertial frames.

[11] Important Note: The superscripts 0, 1, 2, 3 appearing in Equation (12.11) are not exponents; they are used to label the various components. For instance, the superscript "3" in "x^3" refers to coordinate 3 (the z-component in this case), not the cube power of "x", and so on.

Table 12.1: Some quantities in 3-D space and their corresponding 4-D analogs.

Quantity	3-dimensional space	4-dimensional spacetime
Displacement	$\vec{r} = (x,\ y,\ z)$	$x^\mu = (x^0,\ x^1,\ x^2,\ x^3) = (ct, x, y, z) \equiv (x^0,\ \vec{r})$
Gradient operator	$\vec{\nabla} = \left(\frac{\partial}{\partial x},\ \frac{\partial}{\partial y},\ \frac{\partial}{\partial z}\right)$	$\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\partial^0,\ -\vec{\nabla}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t},\ -\vec{\nabla}\right)$
Momentum operator	$\hat{\vec{p}} = (\hat{p}_x,\ \hat{p}_y,\ \hat{p}_z) = -i\hbar\vec{\nabla}$	$\hat{p}^\mu = i\hbar\partial^\mu = \left(p^0,\ \hat{\vec{p}}\right) = \left(\frac{i\hbar}{c}\frac{\partial}{\partial t},\ -i\hbar\vec{\nabla}\right)$
Wave vector	$\vec{k} = \frac{\vec{p}}{\hbar}$	$k^\mu = \frac{1}{\hbar}p^\mu = \left(k^0,\ \vec{k}\right) = \left(\frac{E}{\hbar c},\ \frac{\vec{p}}{\hbar}\right) = \left(\frac{\omega}{c},\ \vec{k}\right)$
Current density	$\vec{j} = (j_x,\ j_y,\ j_z)$	$j^\mu = \left(j^0,\ \vec{j}\right) = \left(c\rho,\ \vec{j}\right)$
Vector potential	$\vec{A} = (A_x,\ A_y,\ A_z)$	$A^\mu = \left(A^0,\ \vec{A}\right) = \left(U/c,\ \vec{A}\right)$
Scalar product	$\vec{A}\cdot\vec{B} = A_xB_x + A_yB_y + A_zB_z$	$A \cdot B = A_\mu B^\mu = A^\mu B_\mu = A^0B^0 - \vec{A}\cdot\vec{B}$
Laplacian & d'Alembertian	$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$	$\Box = \partial_\mu\partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$

summary of the most important 4-vectors and 4-operators[12] along with their 3-dimensional counterparts that are encountered in the study of the Klein–Gordon and Dirac equations.

By convention, we will use the Latin indices $i,\ j,\ k, \cdots$ to refer to the 3-dimensional Euclidean space (hence their ranges are between 1 and 3) and the Greek indices $\mu,\ \nu,\ \alpha, \cdots$ to refer to the 4-dimensional spacetime (hence their ranges are between 0 to and 3).

Contravariant and Covariant notations

In the study of the Klein–Gordon and Dirac equations, we will encounter *upper and lower index* vectors. The **upper index** vector x^μ introduced in (12.11) is known as the **contravariant** form of the 4-vector x; its **covariant** form, which is denoted by the **lower index** vector x_μ, is obtained by merely changing the sign of the spatial components:

$$\boxed{\textbf{Contravariant 4-vector:} \quad x^\mu = (x^0,\ x^1,\ x^2,\ x^3) = (ct,\ x,\ y,\ z) \equiv (x^0,\ \vec{r}),} \tag{12.12}$$

$$\boxed{\textbf{Covariant 4-vector:} \quad x_\mu = (x_0,\ x_1,\ x_2,\ x_3) = (ct,\ -x,\ -y,\ -z) \equiv (x^0,\ -\vec{r}).} \tag{12.13}$$

So, covariant vectors are just spatially inverted contravariant vectors:

$$x^0 = x_0, \qquad x^j = -x_j, \qquad \text{where} \qquad j = 1, 2, 3. \tag{12.14}$$

12.2.4 Minkowski Metric

By analogy to (12.3), we an express any 4-vector A^μ in terms of the four orthonormal basis vectors $(\hat{e}_0,\ \hat{e}_1,\ \hat{e}_2,\ \hat{e}_3)$ of the Minkowski spacetime as follows:

$$A^\mu = A^0\hat{e}_0 + A^1\hat{e}_1 + A^2\hat{e}_2 + A^3\hat{e}_3 = \sum_{\nu=0}^{3} A^\nu\hat{e}_\nu \equiv A^\nu\hat{e}_\nu, \tag{12.15}$$

[12] See Appendix E for their derivation.

where the Einstein summation convention is used in the last step of the previous equation.

An essential feature of special relativity is the invariance or universal constancy of the speed of light in any reference frame; this property is illustrated mathematically by Eq. (12.10). Due to the minus signs in the spatial components, this property defines for us the inner product in the Minkowski spacetime:

$$g(A, A) = A\cdot A = (A^0)^2-(A^1)^2-(A^2)^2-(A^3)^2 \quad \longrightarrow \quad g(A, B) = A\cdot B = A^0B^0-A^1B^1-A^2B^2-A^3B^3, \tag{12.16}$$

where g is the Minkowski metric. Applying (12.16) to $(\hat{e}_0, \hat{e}_1, \hat{e}_2, \hat{e}_3)$, we obtain

$$g(\hat{e}_\mu, \hat{e}_\nu) = \hat{e}_\mu \cdot \hat{e}_\nu = g_{\mu\nu} = \begin{cases} 1, & \mu = \nu = 0 \\ -1, & \mu = \nu = 1,2,3 \\ 0, & \mu \neq \nu \end{cases}, \qquad \mu, \nu = 0,1,2,3 \tag{12.17}$$

where $g_{\mu\nu}$ is the *Minkowski metric*, which can be written in matrix form as follows

$$[g] = (g_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{12.18}$$

In Appendix D, we offer another method for obtaining $[g]$. We should note that relation (12.17) is known as the *Lorentz orthonormality condition*. The structure of the metric (12.18) is a direct consequence of the invariance of the expression (12.10) under the Lorentz transformation (12.9). Using (12.9), we can easily show that the transformed basis vectors $(\hat{e}'_0, \hat{e}'_1, \hat{e}'_2, \hat{e}'_3)$ have the same inner product as $(\hat{e}_0, \hat{e}_1, \hat{e}_2, \hat{e}_3)$:

$$\hat{e}'_\mu \cdot \hat{e}'_\nu = g_{\mu\nu} = \hat{e}_\mu \cdot \hat{e}_\nu, \qquad \mu, \nu = 0,1,2,3. \tag{12.19}$$

Additionally, we can rewrite (12.10) in terms of the metric as follows:

$$x'^\mu g_{\mu\nu} x'^\nu = x^\mu g_{\mu\nu} x^\nu \quad \longrightarrow \quad x'^\mu x'_\mu = x^\mu x_\mu. \tag{12.20}$$

Using the Minkowski metric $g_{\mu\nu}$ and its inverse $g^{\mu\nu}$, we can easily convert contravariant vectors into covariant vectors and vice versa:

$$\boxed{x_\mu = g_{\mu\nu}x^\nu, \qquad x^\mu = g^{\mu\nu}x_\nu,} \tag{12.21}$$

where $g_{\mu\nu}$ is called the covariant metric, and $g^{\mu\nu}$ the contravariant metric. As a result, we can write the inner product between any two 4-vectors in the following more condensed form:

$$A_\mu B^\mu = g_{\mu\nu}A^\mu B^\nu = g^{\mu\nu}A_\mu B_\nu = A^0B^0 - A^iB^i, \qquad \mu,\nu = 0,1,2,3, \quad i,j = 1,2,3. \tag{12.22}$$

We should note that we can interpret $x^\mu x_\mu = c^2t^2 - \vec{r}^2$ as the square of the length of $|x^\mu|$ or, simply, the square of a distance in the Minkowski spacetime. The key feature about distance in the Minkowski space is that it is not going to be necessarily positive. Due to the minus signs in the metric (+, −, −, −), the square of a 4-vector can be positive, zero, or negative. A 4-vector x^μ is called *timelike*, *lightlike*, or *spacelike* if its square is positive ($x^\mu x_\mu > 0$), zero ($x^\mu x_\mu = 0$), or negative ($x^\mu x_\mu < 0$), respectively.

The properties of the Minkowski metric as well as those of 4-vectors are discussed in Appendix E.

Example 12.1
Consider the following 4-vectors in Cartesian coordinates in the Minkowski space time: $A^\mu = (1,-2,0,1)$, $B^\mu = (-4,1,2,-3)$, $C^\mu = (-\sqrt{5},1,2,0)$.
(a) Find the covariant forms of the three 4-vectors listed above.
(b) Find the inner products $A^\mu B_\mu$ and $A^\mu C_\mu$.
(c) Find the causal character of each of the 4-vectors listed above.

Solution
(a) The covariant forms of A^μ, B^μ, C^μ can be obtained by simply changing the signs of their spatial components and leaving the signs of the time components intact:

$$A_\mu = (1,2,0,-1), \qquad B_\mu = (-4,-1,-2,3), \qquad C_\mu = (-\sqrt{5},-1,-2,0). \tag{12.23}$$

(b) Using the general inner product relation: $A^\mu B_\mu = A^0B^0 - A^1B^1 - A^2B^2 - A^3B^3$, we have

$$A^\mu B_\mu = (1)(-4) - (-2)(1) - (0)(2) - (1)(-3) = -3, \tag{12.24}$$

$$A^\mu C_\mu = (1)(-\sqrt{5}) - (-2)(1) - (0)(2) - (1)(0) = -\sqrt{5} - 3, \tag{12.25}$$

(c) Using the general relation for the square of a 4-vector $A^\mu A_\mu = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2$, we obtain

$$A^\mu A_\mu = (1)(1) - (-2)(-2) - (0)(1) - (1)(1) = -4 \implies A^\mu \text{ is spacelike; i.e., non-causal,} \tag{12.26}$$

$$B^\mu B_\mu = (-4)(-4) - (1)(1) - (2)(2) - (-3)(-3) = 2 \implies B^\mu \text{ is timelike; i.e., causal,} \tag{12.27}$$

$$C^\mu C_\mu = (-\sqrt{5})(-\sqrt{5}) - (1)(1) - (2)(2) - (0)(0) = 0 \implies C^\mu \text{ is null.} \tag{12.28}$$

Example 12.2
(a) Show that the Lorentz transformation (12.9) conserves the inner product between any two 4-vectors; i.e., show that $A'^\mu B'_\mu = A^\mu B_\mu$, where A'^μ and B'_μ are boosts of A^μ and B_μ with respect to the first component of spacetime.
(b) Show that the quantity $A^\mu B^\mu = A^0B^0 + A^iB^i$ is not Lorentz invariant; i.e., show that $A'^\mu B'^\mu \neq A^\mu B^\mu$.

Solution
(a) By analogy to (12.9), and since A'^μ and B'_μ are boosts of A^μ and B_μ with respect to the first component, we can write

$$\begin{pmatrix} A'^0 \\ A'^1 \\ A'^2 \\ A'^3 \end{pmatrix} = \begin{pmatrix} \gamma(A^0 - \beta A^1) \\ \gamma(-\beta A^0 + A^1) \\ A^2 \\ A^3 \end{pmatrix}, \qquad \begin{pmatrix} B'^0 \\ B'^1 \\ B'^2 \\ B'^3 \end{pmatrix} = \begin{pmatrix} \gamma(B^0 - \beta B^1) \\ \gamma(-\beta B^0 + B^1) \\ B^2 \\ B^3 \end{pmatrix}, \tag{12.29}$$

which, in turn, lead to

$$A'^0B'^0 = \gamma^2\left(A^0 - \beta A^1\right)\left(B^0 - \beta B^1\right) = \gamma^2\left(A^0B^0 - \beta A^0B^1 - \beta A^1B^0 + \beta^2 A^1B^1\right), \tag{12.30}$$

$$A'^1B'^1 = \gamma^2\left(-\beta A^0 + A^1\right)\left(-\beta B^0 + B^1\right) = \gamma^2\left(\beta^2 A^0B^0 - \beta A^0B^1 - \beta A^1B^0 + A^1B^1\right), \tag{12.31}$$

$$A'^2B'^2 = A^2B^2 \tag{12.32}$$

$$A'^3B'^3 = A^3B^3. \tag{12.33}$$

Subtracting (12.31) from (12.30) and simplifying, we obtain

$$\begin{aligned} A'^0B'^0 - A'^1B'^1 &= \gamma^2\left(1-\beta^2\right)A^0B^0 - \gamma^2\left(1-\beta^2\right)A^1B^1 = \gamma^2\left(1-\beta^2\right)\left(A^0B^0 - A^1B^1\right) \\ &= A^0B^0 - A^1B^1, \end{aligned} \tag{12.34}$$

where we have used the relation $\gamma^2(1-\beta^2) = 1$ in the last step. Now, subtracting (12.32) and (12.33) from (12.34), we obtain

$$A'^0B'^0 - \left(A'^1B'^1 + A'^2B'^2 + A'^3B'^3\right) = A^0B^0 - \left(A^1B^1 + A^2B^2 + A^3B^3\right), \tag{12.35}$$

which proves the Lorentz invariance of the inner product between any two 4-vectors:

$$A'^\mu B'_\mu = A^\mu B_\mu \qquad \text{or} \qquad A'^0B'^0 - A'^iB'^i = A^0B^0 - A^iB^i. \tag{12.36}$$

(b) Adding equations (12.30) and (12.31) and simplifying, we obtain

$$\begin{aligned} A'^0B'^0 + A'^1B'^1 &= \gamma^2\left(1+\beta^2\right)\left(A^0B^0 + A^1B^1\right) - 2\gamma^2\beta\left(A^0B^1 + A^1B^0\right) \\ &= \frac{1+\beta^2}{1-\beta^2}\left(A^0B^0 + A^1B^1\right) - \frac{2\beta}{1-\beta^2}\left(A^0B^1 + A^1B^0\right), \end{aligned} \tag{12.37}$$

where we have used the relation $\gamma^2 = 1/(1-\beta^2)$ in the last step. Now, adding (12.32) and (12.33) to (12.37), we obtain

$$A'^0B'^0 + A'^1B'^1 + A'^2B'^2 + A'^3B'^3 = \frac{1+\beta^2}{1-\beta^2}\left(A^0B^0 + A^1B^1\right) - \frac{2\beta}{1-\beta^2}\left(A^0B^1 + A^1B^0\right) + A^2B^2 + A^3B^3, \tag{12.38}$$

which shows that the Lorentz transformation (12.9) does not conserve the quantity $A'^\mu B'^\mu$:

$$A'^\mu B'^\mu \neq A^\mu B^\mu \qquad \text{or} \qquad A'^0B'^0 + A'^iB'^i \neq A^0B^0 + A^iB^i. \tag{12.39}$$

12.3 Klein–Gordon Equation

12.3.1 Historical note about the Genesis of the Klein–Gordon Equation

Schrödinger's early derivation of what later became known as the Klein–Gordon equation:

> From a historical perspective, Schrödinger was the first who, based on the relativistic energy relation $E^2 = \vec{p}^{\,2}c^2 + m^2c^4$, has introduced an equation that later became known as the Klein–Gordon equation. After deriving this equation, Schrödinger found out that it leads to *negative energy solutions* for which he had no satisfactory explanation; as such, he never published this work and set this equation aside only to return to it later on. But this time, abandoning $E^2 = \vec{p}^{\,2}c^2 + m^2c^4$ and inspired by de Broglie's matter-waves postulate[13], he has started from the nonrelativistic energy relation $E = P^2/(2m)$ for a free particle. Based on de Broglie's idea that a plane

[13] L. de Broglie, *Ann. de Physique*, 3, 22 (1925).

wave, $\exp i(\vec{k}\cdot\vec{r}-\omega t)$ where its angular frequency ω and wave-vector $\vec{k}$ are related by Einstein's relations $E=\hbar\omega$ and $\vec{p}=\hbar\vec{k}$, can be associated with a free particle of mass m, momentum p, and energy E, Schrödinger managed to find a wave equation for it: $i\hbar\partial\psi/\partial t=(-\hbar^2\nabla^2/(2m))\psi$. He then extended the free-particle equation to a particle in a potential V by starting from $E=P^2/(2m)+V$; this work has led to today's well-known Schrödinger equation, $i\hbar\partial\psi/\partial t=(-\hbar^2\nabla^2/(2m)+\hat{V})\psi$, which is the basis of nonrelativistic quantum mechanics and, perhaps, the most important physics equation of the twentieth century. Schrödinger[14] has then successfully applied this equation to a host of problems (step potential, harmonic oscillator, hydrogen atom, etc). Shortly after Schrödinger published his nonrelativistic equation and in their quest of finding a relativistic extension to this equation, yet still unaware about Schrödinger's earlier attempt with the relativistic equation, Gordon (1926) and Klein (1927) have independently introduced their equation by starting from $E^2=\vec{p}^{\,2}c^2+m^2c^4$; their equation was exactly the same as the one introduced by Schrödinger earlier; but, as mentioned above, he had missed the opportunity to publish it before Klein and Gordon. In sum, Schrödinger's earlier unsuccessful attempt with a relativistic quantum relation has paradoxically led to his historical success in laying down the foundations of nonrelativistic quantum mechanics that has inspired generations of physicists (such as Klein, Gordon, Dirac and many others) to emulate his approach to open new fields with far-reaching applications to various areas of physics (such as atomic, nuclear, particle, and solid state physics), chemistry, engineering and, now, to the fast growing field of nanoscience-nanotechology with its systemic applications.

Brief reminder about the nonrelativistic Schrödinger equation:
To obtain the equation of motion of a classical particle of mass m with an energy less than its rest-mass energy, one only needs to substitute the linear momentum, $\vec{p}$, and energy, E, by the respective operators,

$$\vec{p}\longrightarrow -i\hbar\vec{\nabla},\qquad E\longrightarrow i\hbar\frac{\partial}{\partial t}, \tag{12.40}$$

into the **nonrelativistic energy–momentum expression** $E=P^2/2m+V$; this leads to the standard Schrödinger equation

$$i\hbar\frac{\partial\psi(\vec{r},\,t)}{\partial t}=\left[-\frac{\hbar^2}{2m}\nabla^2+\hat{V}(\vec{r})\right]\psi(\vec{r},\,t). \tag{12.41}$$

This equation, which has been the central focus of the first 11 chapters of this text, is invariant under Galilean transformations, but not covariant under Lorentz transformations due to the asymmetry between the space and time derivatives. Since the space and time derivatives are of different order (second order in space and first order in time), Eq (12.41) changes its structure under transformations from one inertial system to another.

To prepare the stage for the interpretation of the conservation of probability within the context of the Klein–Gordon equation, it is worth briefly recalling the derivation of the probability density, the current density and the continuity equation we have presented in Chapter 3

[14]E. Schrödinger, *Ann. Phys.*, 79, 361 (1926); 79, 489 (1926); 79, 734 (1926); 80, 437 (1926); 81, 109 (1926); *Die Naturwissenschaften*, 14, 664 (1926); *Phys. Rev.*, Vol. 28, No. 6, 1049 (1926); *Collected Papers on Wave Mechanics*, Blackie & Son, London, 1928.

(Subsection 3.6.4) for the Schrödinger equation. Multiplying (12.41) on the left by $\psi^*(\vec{r},t)$ and the complex conjugate of (12.41) on the right by $\psi(\vec{r},t)$ and then subtracting the two resulting equations (i.e., $\psi^*(\vec{r},t)\times(12.41) - (12.41)^* \times \psi(\vec{r},t)$), we obtain

$$i\hbar\frac{\partial}{\partial t}\left[\psi^*(\vec{r},t)\psi(\vec{r},t)\right] = -\frac{\hbar^2}{2m}\left[\psi^*(\vec{r},t)\nabla^2\psi(\vec{r},t) - \left(\nabla^2\psi^*(\vec{r},t)\right)\psi(\vec{r},t)\right]. \tag{12.42}$$

This equation can be rewritten in the following well known form – the continuity equation:

$$\frac{\partial\rho(\vec{r},t)}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0, \tag{12.43}$$

where $\rho(\vec{r},t)$, the probability density, and $\vec{J}$, the current density, are given by

$$\rho(\vec{r},t) = \psi^*(\vec{r},t)\psi(\vec{r},t) = \left|\psi(\vec{r},t)\right|^2, \qquad \vec{J}(\vec{r},t) = \frac{i\hbar}{2m}\left(\psi\vec{\nabla}\psi^* - \psi^*\vec{\nabla}\psi\right). \tag{12.44}$$

As seen in Chapter 3, the probability density $\rho(\vec{r},t) = \left|\psi(\vec{r},t)\right|^2$ is *positive definite*. As we shall see in Subsection 12.3.4.2 (see Equation (12.83)), the density is not necessarily positive within the context of the Klein–Gordon equation; this turned out to be one of the most untractable problems of the Klein–Gordon equation.

12.3.2 Klein–Gordon Equation for a Free particle and its Covariance

The Klein–Gordon approach deals with the description of the motion of a relativistic *single particle*. From the very outset, we should mention that the single-particle nature of the Klein–Gordon equation leads to a number of untractable problems (see Subsection 12.3.2) and paradoxes (see Subsection 12.3.5) that cannot be surmounted until we treat it as a field equation describing fields (i.e., multi-particle systems), not single particles. But for pedagogical reasons, we are going to follow the standard historical approach where we treat the Klein–Gordon equation as describing a single particle, discuss the problems and anomalies generated by this approach, and then show that the only way out is by treating it as a field equation.

To illustrate the essential elements of the Klein–Gordon equation and to simplify the discussion, let us consider the simple physical system of an isolated free particle with an energy appreciably higher than its rest-mass energy (i.e., a relativistic free particle). The relativistic relationship between its energy E, its linear momentum $\vec{p}$, and its rest-mass energy mc^2 is provided by the theory of special relativity:

$$E = \sqrt{\vec{p}^{\,2}c^2 + m^2c^4}. \tag{12.45}$$

To find a quantum mechanical description of a free single particle, one needs to quantify (12.45) by simply using the operator expressions corresponding to E and $\vec{p}^{\,2}$ (i.e., $E \longrightarrow -\hbar\partial/\partial t$ and $\vec{p}^{\,2} \longrightarrow -\hbar^2\vec{\nabla}^2$) which are listed in (12.40) and applying them on the particle's wave function $\psi(\vec{r},\,t)$; this leads to

$$\hbar\frac{\partial\psi(\vec{r},\,t)}{\partial t} = \sqrt{m^2c^4 - \hbar^2c^2\nabla^2}\,\psi(\vec{r},\,t). \tag{12.46}$$

This equation suffers from a major disadvantage in that it is not covariant due to the asymmetry between time and space as one would expect for a relativistic theory; as will be discussed below, covariance is an essential requirement in the equation we are going to utilize. The presence of the square root in (12.46) makes it hard to interpret. It helps to mention that one might consider expanding the expression under the square-root sign of (12.46) as a power series in $\vec{\nabla}^2$, but the resulting series may not always converge. Hence, this route in an attempt to find a suitable relativistic equation is forcibly going to be fraught with a number of unnecessary technical as well as conceptual complications.

For the sake of avoiding the formalism's complications that arise from the square-rooted energy expression (12.46), one needs to look for a simpler energy–momentum relation. The alternative route consists of using an energy–momentum expression which is readily covariant. To avoid the non-covariance problems inherent in Eq (12.45), Klein and Gordon have started with E^2, not E; that is, they started from the following relativistic energy–momentum relation

$$E^2 = \vec{p}^{\,2}c^2 + m^2c^4 \tag{12.47}$$

and then quantified it by using the operator expressions corresponding to E^2 and $\vec{p}^{\,2}$: $E^2 \longrightarrow (i\hbar\partial/\partial t)^2$ and $\vec{p}^{\,2} \longrightarrow \left(-i\hbar\vec{\nabla}\right)^2$. This leads to

$$\left(i\hbar\frac{\partial}{\partial t}\right)^2 \psi(\vec{r},\, t) = c^2\left(-i\hbar\nabla\right)^2 \psi(\vec{r},\, t) + m^2c^4\psi(\vec{r},\, t), \tag{12.48}$$

or to

$$-\hbar^2\frac{\partial^2\psi(\vec{r},\, t)}{\partial t^2} = -\hbar^2c^2\nabla^2\psi(\vec{r},\, t) + m^2c^4\psi(\vec{r},\, t); \tag{12.49}$$

which can be rewritten as

$$\boxed{\frac{1}{c^2}\frac{\partial^2\psi(\vec{r},\, t)}{\partial t^2} = \left(\nabla^2 - \frac{m^2c^2}{\hbar^2}\right)\psi(\vec{r},\, t).} \tag{12.50}$$

This is known as the Klein–Gordon equation for a free particle. This equation can abbreviated into the following covariant form

$$\boxed{\left(\partial_\mu\partial^\mu + \frac{m^2c^2}{\hbar^2}\right)\psi(x^\mu) = 0 \quad \longrightarrow \quad \left(\Box + \frac{m^2c^2}{\hbar^2}\right)\psi(x^\mu) = 0,} \tag{12.51}$$

where $x^\mu = (ct,\, \vec{r})$ and $\Box = \partial_\mu\partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$ is the d'Alembertian or wave operator which was introduced in Table 12.1; it is worth recalling that that $\hbar/mc$ is the Compton wavelength. Unlike (12.46), equation (12.51) is covariant as will be proven below. Additionally, it is worth mentioning at this level that, unlike the Schrödinger equation which lacks the symmetry between time and space (first order in time and second order in space), the Klein–Gordon equation has the desired symmetry between space and time; it treats time and space on an equal footing (second order both in time and space).

12.3.2.1 Lorentz Invariance of the Klein–Gordon Equation

As mentioned above, relativistic covariance is easier to see when one is using the covariance notation. The Klein–Gordon equation (12.51) is manifestly covariant; it keeps the same form in all inertial frames. Due to the symmetry of the space and time derivatives (second order both in space and time), the Klein–Gordon equation will have the same structure as we change from one inertial frame to another. At this level, it helps to treat the terms "invariance" and "covariance" with some caution as they cannot be used interchangeably and they don't mean the same thing. On the one hand, a quantity is said to be "invariant" if it does not change from one frame of reference to another; this applies, for instance, to scalars such as inner products (such as the d'Alembertian) or to constants (such as the speed of light or terms like $m^2c^2/\hbar^2$ present in (12.51)). On the other hand, an expression (or relation) is said to be "covariant" if it keeps the same form or structure under a Lorentz transformation; this applies to physical laws or to covariant relations such as the Klein–Gordon equation or the Dirac equation as we will see below.

The Klein–Gordon equation is not merely covariant, it is Lorentz invariant as well. Under an arbitrary Lorentz transformation $x^\mu \longrightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu$, the Klein–Gordon equation (12.51) transforms like

$$\left(\Box + \frac{m^2c^2}{\hbar^2}\right)\psi(x^\mu) = 0 \qquad \Longleftrightarrow \qquad \left(\Box' + \frac{m^2c^2}{\hbar^2}\right)\psi'(x'^\mu) = 0. \tag{12.52}$$

But, since the term $m^2c^2/\hbar^2$ is a constant, the d'Alembertian $\Box = \partial_\mu\partial^\mu$ is a scalar quantity (hence, $\Box' = \Box$), and ψ is a *scalar function*[15]

$$\psi'(x'^\mu) \equiv \psi'(\Lambda^\mu{}_\nu x^\nu) = \psi(x^\mu), \tag{12.53}$$

we have

$$\left(\Box' + \frac{m^2c^2}{\hbar^2}\right)\psi'(x'^\mu) = \left(\Box + \frac{m^2c^2}{\hbar^2}\right)\psi(x^\mu) = 0, \tag{12.54}$$

and, hence, *the Klein–Gordon equation is Lorentz covariant*; it has the same form in all reference frames. Now, since $\psi'(x'^\mu) = \psi(x^\mu)$, all directions in spacetime are equivalent (no orientation is privileged); that is, ψ is invariant under rotations. Hence, *the Klein–Gordon equation must describe* spin-zero particles.

12.3.3 Nonrelativistic Limit of the Klein–Gordon Equation

To gain additional insight into the Klein–Gordon equation, most notably its connection with the Schrödinger equation, let us study its nonrelativistic limit. In this limit, the particle's rest-mass energy mc^2 is much greater than all energy terms; for instance, in the nonrelativistic limit, the kinetic energy $p^2/2m$ is going to be too small compared to mc^2: $p^2/2m \ll mc^2$. For this, let us write the wave function $\psi(x^\mu)$ as a product of two factors: one factor $\phi(x^\mu)$ that involves the particle's kinetic energy and the other factor $e^{-imc^2t/\hbar}$ that describes oscillations due to the particle's rest-mass energy mc^2:

$$\psi(x^\mu) = \phi(x^\mu)e^{-imc^2t/\hbar}. \tag{12.55}$$

[15]The wave function is a complex-valued function, also notice that $\psi(x^\mu)$ and $\psi'(x'^\mu)$ refer to the same spacetime point. Being a relativistic scalar, the wave function ψ is nothing but a complex number that has the same value in all reference frames.

In what follows, we are going to focus on finding an equation for $\phi(x^\mu)$ and will show that it is similar to the Schrödinger equation. First, let us take the first time derivative of (12.55):

$$\frac{\partial \psi}{\partial t} = \left(\frac{\partial \phi}{\partial t} - \frac{imc^2}{\hbar}\phi\right) e^{-imc^2 t/\hbar}. \tag{12.56}$$

Second, since

$$\frac{\partial}{\partial t}\left(\frac{\partial \phi}{\partial t} - \frac{imc^2}{\hbar}\phi\right) = \frac{\partial^2 \phi}{\partial t^2} - \frac{imc^2}{\hbar}\frac{\partial \phi}{\partial t}, \tag{12.57}$$

we can calculate the second time derivative of (12.55) as follows

$$\begin{aligned}
\frac{\partial^2 \psi}{\partial t^2} &= \left[\frac{\partial}{\partial t}\left(\frac{\partial \phi}{\partial t} - \frac{imc^2}{\hbar}\phi\right)\right] e^{-imc^2 t/\hbar} + \left(\frac{\partial \phi}{\partial t} - \frac{imc^2}{\hbar}\phi\right)\frac{\partial}{\partial t} e^{-imc^2 t/\hbar} \\
&= \left(\frac{\partial^2 \phi}{\partial t^2} - \frac{imc^2}{\hbar}\frac{\partial \phi}{\partial t} - \frac{imc^2}{\hbar}\frac{\partial \phi}{\partial t} - \frac{im^2c^4}{\hbar^2}\phi\right) e^{-imc^2 t/\hbar} \\
&= \left(\frac{\partial^2 \phi}{\partial t^2} - \frac{2imc^2}{\hbar}\frac{\partial \phi}{\partial t} - \frac{m^2c^4}{\hbar^2}\phi\right) e^{-imc^2 t/\hbar}.
\end{aligned} \tag{12.58}$$

Now, applying the d'Alembertian operator $\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$ on (12.55) and using (12.58), we have

$$\begin{aligned}
\Box\psi(x^\mu) &= \frac{1}{c^2}\frac{\partial^2 \psi}{\partial t^2} - \nabla^2\psi \\
&= \left(\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} - \frac{2im}{\hbar}\frac{\partial \phi}{\partial t} - \frac{m^2c^2}{\hbar^2}\phi\right) e^{-imc^2 t/\hbar} - \left(\nabla^2\phi\right) e^{-imc^2 t/\hbar} \\
&= \left(\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} - \frac{2im}{\hbar}\frac{\partial \phi}{\partial t} - \frac{m^2c^2}{\hbar^2}\phi - \nabla^2\phi\right) e^{-imc^2 t/\hbar}.
\end{aligned} \tag{12.59}$$

Now, inserting (12.59) into (12.51), we obtain

$$\left(\Box + \frac{m^2c^2}{\hbar^2}\right)\psi(x^\mu) = \left(\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} - \frac{2im}{\hbar}\frac{\partial \phi}{\partial t} - \frac{m^2c^2}{\hbar^2}\phi - \nabla^2\phi\right) e^{-imc^2 t/\hbar} + \frac{m^2c^2}{\hbar^2}\phi e^{-imc^2 t/\hbar} = 0, \tag{12.60}$$

which, after canceling terms, leads to

$$\left(\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} - \frac{2im}{\hbar}\frac{\partial \phi}{\partial t} - \nabla^2\phi\right) e^{-imc^2 t/\hbar} = 0 \quad \Longrightarrow \quad \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} - \frac{2im}{\hbar}\frac{\partial \phi}{\partial t} - \nabla^2\phi = 0. \tag{12.61}$$

Now, let us focus on the nonrelativistic limit of the Klein–Gordon equation. In the nonrelativistic limit, the speeds of the particles are too small compared to the speed of light. So, by taking the limit $c \longrightarrow \infty$ in (12.61), the term $\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2}$ can be neglected and, hence, we end with the following equation

$$-\frac{2im}{\hbar}\frac{\partial \phi}{\partial t} - \nabla^2\phi = 0. \tag{12.62}$$

After multiplying both sides by $\hbar^2/2m$, this equation leads at once to

$$i\hbar\frac{\partial \phi(x^\mu)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\phi(x^\mu). \tag{12.63}$$

This is nothing other than the Schrödinger equation for a *spinless* free particle. This is why the Klein–Gordon equation is said to describe *spin-zero* particles. Additionally, the Klein–Gordon equation can be viewed as the direct relativistic generalization of the Schrödinger equation.

12.3.4 Conceptual Problems of the Klein–Gordon Equation

The Klein–Gordon equation (12.51), albeit satisfying the laws of relativity, suffers from several fundamental problems which we are going to discuss below.

12.3.4.1 First Problem: Negative Energies

The Klein–Gordon equation (12.51) for a free particle is a differential equation with constant coefficients and, hence, its solutions are plane-waves,

$$\psi(x^\mu) = Ne^{-i(Et-\vec{p}\cdot\vec{r})/\hbar} = Ne^{-ip_\mu x^\mu/\hbar}, \tag{12.64}$$

where N is a normalization constant. It helps noting that, since the inner product $p_\mu x^\mu$ and N in (12.64) are scalars, *the wave function $\psi(x^\mu)$ transforms like a scalar as well under Lorentz transformations*. This is expected since, as will be seen below, the Klein–Gordon equation describes spin-zero particles and, hence, the wave function has only one component very much like the solutions of the nonrelativistic Schrödinger equation for spin-zero particles.

Applying $\partial_\mu\partial^\mu$ on (12.64), we obtain[16]

$$\begin{aligned}\partial_\mu\partial^\mu\psi(x^\mu) &= N\frac{\partial}{\partial x_\mu}\left(\frac{\partial}{\partial x^\mu}e^{-ip_\mu x^\mu/\hbar}\right) = N\frac{\partial}{\partial x_\mu}\left(-\frac{i}{\hbar}p_\mu e^{-ip^\mu x_\mu/\hbar}\right) \equiv -\frac{1}{\hbar^2}p_\mu p^\mu\left(Ne^{-ip_\mu x^\mu/\hbar}\right)\\ &= -\frac{1}{\hbar^2}p_\mu p^\mu\psi(x^\mu).\end{aligned} \tag{12.65}$$

Substituting (12.65) into (12.51), we end up with we obtain

$$\left(-\frac{1}{\hbar^2}p_\mu p^\mu + \frac{m^2c^2}{\hbar^2}\right)\psi(x^\mu) = 0 \qquad \Longrightarrow \qquad p_\mu p^\mu = m^2c^2, \tag{12.66}$$

which, when combined with $p_\mu p^\mu = E^2/c^2 - \vec{p}^{\,2}$, leads to

$$\frac{E^2}{c^2} - \vec{p}^{\,2} = m^2c^2 \qquad \Longrightarrow \qquad E^2 = \vec{p}^{\,2}c^2 + m^2c^4, \tag{12.67}$$

and, hence, to

$$E = \pm\sqrt{\vec{p}^{\,2}c^2 + m^2c^4}. \tag{12.68}$$

This clearly indicates that the Klein–Gordon equation yields positive as well as *negative energy* solutions[17]. This was expected, since, for the sake of dealing with a simple equation and avoiding the formalism's complications arising from the square-rooted energy expression $E = \sqrt{\vec{p}^{\,2}c^2 + m^2c^4}$, by starting with the squared energy expression $E^2 = \vec{p}^{\,2}c^2 + m^2c^4$ one was

[16] In the following equation, we have used the fact that $p_\mu x^\mu \equiv p^\mu x_\mu$ in the exponent; i.e., $e^{-ip_\mu x^\mu/\hbar} \equiv e^{-ip^\mu x_\mu/\hbar}$.

[17] As mentioned in the Historical Note above, when Schrödinger obtained this negative energy solution, he abandoned this relativistic equation and developed the nonrelativistic Schrödinger equation instead.

bound to introduce extraneous negative-energy solutions, $E = -\sqrt{\vec{p}^{\,2}c^2 + m^2c^4}$. So the Klein-Gordon equation (12.51) has solutions that consist of two energy continua $-\infty < E < -mc^2$ and $+mc^2 < E < +\infty$ and that are separated by a forbidden region of width $2mc^2$; clearly, the particle has no ground state. Actually, this energy spectrum is similar to that of a free Dirac particle as illustrated in Figure 12.2 below.

While the positive-energy solutions, $E = \sqrt{\vec{p}^{\,2}c^2 + m^2c^4}$, are physically meaningful, the negative-energy solutions do not lend themselves to a physically acceptable interpretation within the single-particle Klein–Gordon formalism. One might be tempted to discard them on physical grounds; but one simply cannot omit them in such an arbitrary manner. Thus, the energy expression (12.68), though more general than (12.45), introduces a formidable ambiguity emanating from the negative-energy solutions, for they have no analog either in nonrelativistic quantum mechanics or in the classical theory of relativity. During the development of relativistic quantum mechanics, it took several years to figure out that the negative-energy solutions are associated with the existence of antiparticles. Their existence, much less their prediction, is not possible within the context of the Klein–Gordon equation due to its single-particle nature; they, however, came out naturally from Dirac's equation, as will be explained below.

12.3.4.2 Second Problem: Negative Probability Densities:

First, let us derive the the 4-current density from the Klein–Gordon equation (12.51) and find out if it satisfies the continuity equation in the covariant form. Using the fact that the d'Alembert operator $\Box = \partial_\mu\partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$ is Hermitian,

$$\Box^\dagger = \left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)^\dagger = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 = \Box, \tag{12.69}$$

we can write the complex conjugate of the Klein–Gordon equation (12.51) as follows

$$\left(\Box + \frac{m^2c^2}{\hbar^2}\right)\psi^*(x^\mu) = 0. \tag{12.70}$$

By analogy to the Schrödinger equation approach outlined above (see Eqs (12.42)–(12.44)), multiplying (12.51) on the left by $\psi^*(x^\mu)$ and (12.70) on the right by $\psi(x^\mu)$ then subtracting the two resulting equations (i.e., $\psi^*(x^\mu)\times$(12.51) − (12.70)$\times\psi(x^\mu)$), we obtain[18]

$$\psi^*\left(\Box\psi + \frac{m^2c^2}{\hbar^2}\psi\right) - \left(\Box\psi^* + \frac{m^2c^2}{\hbar^2}\psi^*\right)\psi = 0, \qquad \text{or} \qquad \psi^*\Box\psi - \psi\Box\psi^* = 0; \tag{12.71}$$

since $\Box = \partial_\mu\partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$, we can can rewrite $\psi^*\Box\psi - \psi\Box\psi^* = 0$ as

$$\frac{1}{c^2}\left(\psi^*\frac{\partial^2\psi}{\partial t^2} - \psi\frac{\partial^2\psi^*}{\partial t^2}\right) - \left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right) = 0. \tag{12.72}$$

[18] In deriving the second equation in (12.71) we have used the facts that $\psi^*\psi \equiv \psi\psi^*$ and $\left(\Box\psi^*\right)\psi \equiv \psi\Box\psi^*$.

Using the following two relations

$$\psi^*\frac{\partial^2\psi}{\partial t^2} - \psi\frac{\partial^2\psi^*}{\partial t^2} = \frac{\partial}{\partial t}\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right), \tag{12.73}$$

$$\psi^*\nabla^2\psi - \psi\nabla^2\psi^* = \vec{\nabla}\cdot\left(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right), \tag{12.74}$$

we can reduce (12.72) to

$$\frac{1}{c^2}\frac{\partial}{\partial t}\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right) - \vec{\nabla}\cdot\left(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right) = 0, \tag{12.75}$$

which, in turn, when multiplied by $i\hbar/2m$ (to obtain physically meaningful dimensions for ρ and $\vec{j}$), leads to the *continuity equation*

$$\frac{\partial\rho(x^\mu)}{\partial t} + \vec{\nabla}\cdot\vec{j}(x^\mu) = 0, \tag{12.76}$$

where $\rho(x^\mu)$ and $\vec{j}(x^\mu)$ represent the probability density and the current density, respectively:

$$\boxed{\rho(x^\mu) = \frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right), \qquad \vec{j}(x^\mu) = \frac{\hbar}{2im}\left(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right).} \tag{12.77}$$

It is worth noticing that Equation (12.76) is the 4-dimensional equivalent of the 3-dimensional continuity equation (12.43). In fact, we can rewrite (12.76) in the following covariant form

$$\boxed{\frac{\partial\rho(x^\mu)}{\partial t} + \vec{\nabla}\cdot\vec{j}(x^\mu) = 0 \quad \Longrightarrow \quad \partial_\mu j^\mu = 0, \qquad j^\mu = \left(c\rho(x^\mu),\ \vec{j}(x^\mu)\right),} \tag{12.78}$$

where j^μ is the *4-current*; it can be written as

$$\boxed{j^\mu = \frac{i\hbar}{2m}\left(\psi^*\partial^\mu\psi - \psi\partial^\mu\psi^*\right) = \frac{i\hbar}{2m}\left(\psi^*\frac{\partial\psi}{\partial x_\mu} - \psi\frac{\partial\psi^*}{\partial x_\mu}\right),} \tag{12.79}$$

where, as seen above, ∂^μ is given by $\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial x^0}, -\vec{\nabla}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\vec{\nabla}\right)$.

So far, we have shown that the Klein–Gordon equation yields a density, $\rho(x^\mu)$, and a current density, $\vec{j}(x^\mu)$, that satisfy a continuity equation (12.76); or, in the covariant notation, the Klein–Gordon equation offers a 4-current vector j^μ (see Eq (12.79)) that satisfies a covariant continuity equation $\partial_\mu j^\mu = 0$ (see (12.78)).

On the face of it, everything appears to be fine up to now. However, an important question arises at this level: Can we interpret $\rho(x^\mu)$ as a *probability density*? The answer, as will be clarified below, is a disappointing no; the reason is starkly simple: the negative-energy solutions give a negative value for the density $\rho(x^\mu)$ and, hence, it cannot be a probability. This conclusion can be easily shown by plugging the plane wave solution (12.64) into (12.77); that is, using the following relations

$$\frac{\partial}{\partial t}\psi(x^\mu) = \frac{\partial}{\partial t}\left(Ne^{-i(Et-\vec{p}\cdot\vec{r})/\hbar}\right) = -\frac{i}{\hbar}E\psi(x^\mu), \tag{12.80}$$

$$\frac{\partial}{\partial t}\psi^*(x^\mu) = \frac{\partial}{\partial t}\left(N^*e^{i(Et-\vec{p}\cdot\vec{r})/\hbar}\right) = \frac{i}{\hbar}E\psi^*(x^\mu) \tag{12.81}$$

we obtain[19]

$$\rho(x^\mu) = \frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right) = \frac{i\hbar}{2mc^2}\left(-\frac{i}{\hbar}E\psi^*\psi - \frac{i}{\hbar}E\psi\psi^*\right), \tag{12.82}$$

which, when combined with $\psi = Ne^{-ip_\mu x^\mu/\hbar}$, leads to

$$\boxed{\rho(x^\mu) = \frac{E}{mc^2}|\psi|^2 = |N|^2\frac{E}{mc^2}.} \tag{12.83}$$

This equation shows that the density is proportional to the energy of the particle, E; this leads to a serious problem. Why is this a problem? As seen in (12.68) where $E = \pm\sqrt{\vec{p}^{\,2}c^2 + m^2c^4}$, the Klein–Gordon equation has positive as well as negative energy solutions; *for the case of the negative-energy solutions, the density ρ is negative* which simply makes no sense; it is simply unphysical. This unphysical result rules out the possibility of interpreting ρ as a conventional probability density which must always be positive. Another way to look at it is as follows: as proven above (right after equation (12.64)), since ψ is invariant (a Lorentz scalar), it cannot represent a density, for a density is expected to transform like the time-like component of a 4-vector due to volume contraction under Lorentz transformations. By giving rise to negative densities (or norms), the Klein–Gordon equation violates one of the most sacrosanct tenets of quantum mechanics – the probability density is always positive definite. The negative densities problem was one of the main historical reasons that led to the rejection of the Klein–Gordon equation as a viable theory to describe the motion of single particles.

12.3.4.3 Third Problem: Time-Evolution Indeterminacy of the Wave Function

Unlike the Schrodinger equation which is first-order in time, the Klein–Gordon equation (12.50) is second-order in time. While only one boundary condition is sufficient to find the solutions of the Schröinger equation, we would necessarily need two boundary conditions to determine the solutions of the Klein–Gordon equation; namely, we need to know both $\psi(x^\mu)$ and $\partial\psi(x^\mu)/\partial t$ at $t = 0$ in order to find $\psi(x^\mu)$ for $t > 0$. This can be seen as follows. Using (12.64), and due to the emergence of the two energy values (positive and negative) that have resulted from the second order time derivatives in the Klein–Gordon equation, we can write the most general solution as follows

$$\psi(x^\mu) = N_1e^{-i(E^{(+)}t - \vec{p}\cdot\vec{r})} + N_2e^{-i(E^{(-)}t - \vec{p}\cdot\vec{r})}, \tag{12.84}$$

where

$$E^{(\pm)} = \pm\sqrt{\vec{p}^{\,2}c^2 + m^2c^4}, \tag{12.85}$$

and where N_1 and N_2 are two unknown normalization constants. Clearly, two boundary conditions are needed to determine these two unknown constants. Thus, an extra degree of freedom is introduced in comparison to the Schrödinger equation which, due to its first order dependence on time derivatives, requires only to know ψ at one time in order to determine its value at any later time; this is one of the most essential tenets of quantum mechanics. Manifestly, the Klein–Gordon equation fails to adhere to this basic tenet, for it

[19]Notice that by having multiplied (12.77) by $i\hbar/2m$, the ρ obtained in (12.83) has the correct physical dimensions of a probability density, since $|N|^2$ is the number of particles per unit volume.

lacks a prescription to determine the wave function at later times from its value at $t = 0$. This problem justifies the need to look for a relativistic equation that contains only a first order derivative in time. As will be seen in the next section, this is what has motivated Dirac to search for an equation that involves only first-order time derivatives.

To summarize, here are the main problems encountered by the Klein–Gordon equation:

- Being based on the relativistic momentum-energy expression $E^2 = \vec{p}^{\,2}c^2 + m^2c^4$, the Klein–Gordon equation yields both positive and negative-energy solutions and fails to provide a physically meaningful interpretation for the negative-energy solutions.
- Although the Klein–Gordon equation has a density and a current-density that obey a continuity equation, the density is not positive definite; this constitutes a major departure from quantum mechanics.
- Being second-order in time, the Klein–Gordon equation fails to provide a prescription for determining the wave function at a given time from one of its values at an earlier time.

As these problems turned out to be untractable, it became evident that the Klein–Gordon equation cannot be a viable theory of relativistic quantum mechanics.

The problems outlined above are due essentially to *the single-particle* interpretation of the Klein–Gordon equation. Shortly after its appearance in 1927, this equation fell into disrepute until 1934 when Pauli and Weisskopf[20] managed to resurrect it by showing that it can be reinterpreted as a *field equation* for *spinless* particles (see Subsection 13.5); they used it to describe mesons with spin zero, such as π–mesons, through the use of second quantization. The Klein–Gordon equation works for multi-particle systems, not for single particles; its solutions include *scalar fields* whose quanta are *spin-zero* particles. As such, the Klein–Gordon equation has been serving as the basis for field theories where the negative-energy states can be interpreted as representing antiparticles. The above mentioned problems and internal inconsistencies notwithstanding, the Klein–Gordon equation is accepted today as offering a correct description of relativistic *spin-zero*, or *scalar*[21], particles such as the π-mesons or the recently discovered Higgs boson particle particle.

In what follows, we are going to apply the Klein–Gordon equation to the problem of scattering charged, spin-zero particles from a step potential and show that the equation runs into an anomaly which cannot be resolved until we abandon its *single-particle* nature and posit the creation of particle-antiparticle pairs. This anomaly shows clearly that, due to their single-particle nature, relativistic quantum mechanical theories such as the Klein–Gordon equation are not suitable for describing relativistic phenomena.

12.3.5 Klein Paradox for Spin-Zero Particles

In what follows, we are going to solve the Klein–Gordon equation for the scattering of a relativistic, charged spin-zero particle of energy E from a step potential V_0 as displayed in Figure 12.1. As a reminder, we have solved this problem for a non-relativistic spinless particle

[20] W. Pauli and V. Weisskopf, Helv. Phys. Acta 7, 709 (1934)

[21] *Scalar Particle*: it is a particle that has no spin and no internal structure. The Higgs boson is an example of a scalar particle.

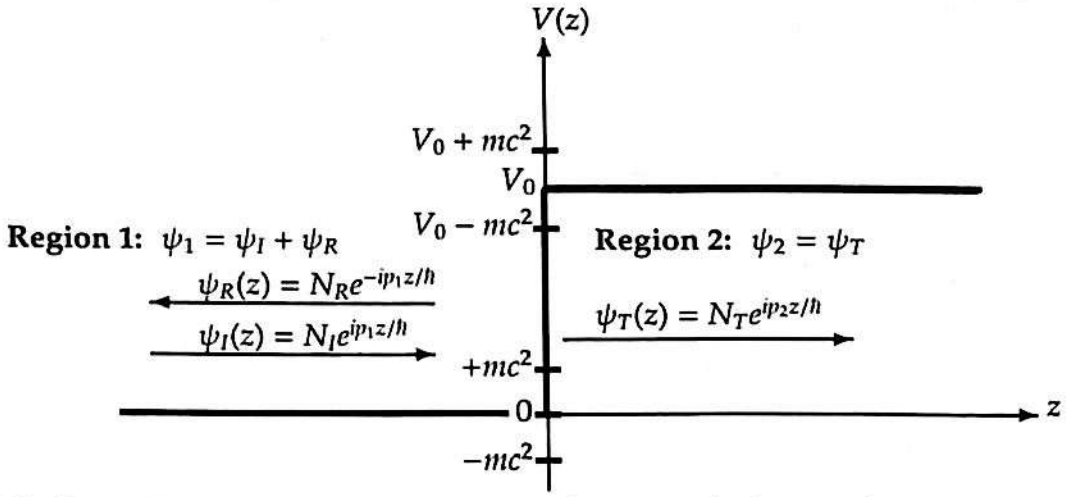


Figure 12.1: Scattering of a free relativistic, spinless particle from a Step potential V_0; the propagation directions of the incident, reflected, and transmitted waves corresponding to the energy range $E > V_0 + mc^2$ are shown in the figure.

in Chapter 4 (see Section 4.4) using, obviously, the Schrödinger equation. The main findings were as follows: when $E > V_0$, we have both reflection and transmission, whereas when $E < V_0$ we have total reflection but with an exponential decay of the wave in the classically forbidden region $z > 0$. In the Klein–Gordon situation, however, the solutions are going to involve some new physics. First, we are going to treat the Klein–Gordon equation as a standard equation of relativistic quantum mechanics where the wave function Ψ describes a single-particle, much like the Schrödinger equation or the approach adopted in Subsection 12.3.2 above, and then examine the outcome resulting from this approach to see if we need to modify something. We are going to see that, when the potential V_0 is strong, the solutions of the Klein–Gordon equation lead to a paradox – the Klein paradox. We will then see how to interpret these solutions and how to resolve this paradox.

As illustrated in Figure 12.1, we would need to solve the Klein–Gordon equation in two regions: (a) Region 1, corresponding to $z < 0$, and (b) Region 2, corresponding to $z \geq 0$. We assume that the particle is moving along the positive z–axis and that it has positive energy in Region 1, $E = \sqrt{c^2 p_1^2 + m^2 c^4}$, where p_1 is the momentum of the particle in Region 1. Since we have a one-dimensional problem where the incident particle is moving along the positive z–axis, we can write the Klein–Gordon equation (12.50) as follows:

$$\left(\frac{d^2}{dz^2} - \frac{m^2c^2}{\hbar^2}\right)\Psi(z,\, t) = \frac{1}{c^2}\frac{\partial^2\Psi(z,\, t)}{\partial t^2}. \tag{12.86}$$

While the particle is free in Region 1, it will be moving under the influence of a constant potential V_0 in Region 2. Since the potential is time-independent, the time and spatial components of the solutions in the two regions will separate:

$$\Psi(z,\, t) = e^{-iEt/\hbar}\psi(z). \tag{12.87}$$

Inserting $\Psi(z,\, t)$ into (12.86) and cancelling the time-dependent exponent, $e^{-iEt/\hbar}$, we obtain

two equations that correspond to regions 1 and 2, respectively, which involve only the time-independent wave function $\psi(z)$:

$$\left(\frac{d^2}{dz^2} - \frac{m^2c^2}{\hbar^2}\right)\psi(z) = -\frac{E^2}{\hbar^2c^2}\psi(z), \qquad (z < 0), \tag{12.88}$$

$$\left(\frac{d^2}{dz^2} - \frac{m^2c^2}{\hbar^2}\right)\psi(z) = -\frac{(E - V_0)^2}{\hbar^2c^2}\psi(z), \qquad (z \geq 0). \tag{12.89}$$

In what follows, we are going to solve the above equations for the following three energy ranges.

Case $E > V_0 + mc^2$: Oscillating reflected and transmitted waves:
As depicted in Figure 12.1, while the most general solution of (12.88) consists of incident and reflected plane waves, the solutions of (12.89) consists of a transmitted plane wave propagating to the right (no wave gets reflected from the $z > 0$ region):

$$\psi(z) = \begin{cases} \psi_1(z) = \psi_I(z) + \psi_R(z) = N_I e^{ip_1 z/\hbar} + N_R e^{-ip_1 z/\hbar} & z < 0 \\ \psi_2(z) = \psi_T(z) = N_T e^{ip_2 z/\hbar} & z \geq 0, \end{cases} \tag{12.90}$$

where N_I, N_R, and N_T are normalization constants and where p_1 and p_2 are the z–components of the particle's momentum in Regions 1 and 2:

$$p_1 = \sqrt{\frac{E^2}{c^2} - m^2c^2}, \qquad p_2 = \pm\sqrt{\frac{E_2^2}{c^2} - m^2c^2} = \pm\sqrt{\frac{(E - V_0)^2}{c^2} - m^2c^2}, \qquad E_2 = E - V_0, \tag{12.91}$$

where E_1 is the particle's energy in Region 1 and E_2 the energy in Region 2; for generality, we have considered that p_2 can be either positive or negative as we will see below. We can write the respective 4-vector momenta corresponding to the incident, reflected, and transmitted waves as follows:

$$p_I^\mu = \left(\frac{E}{c}, 0, 0, p_1\right), \qquad p_R^\mu = \left(\frac{E}{c}, 0, 0, -p_1\right), \qquad p_T^\mu = \left(\frac{E_2}{c}, 0, 0, p_2\right). \tag{12.92}$$

Now, we can obtain the relations between the constants N_I, N_R, and N_T from the boundary conditions. Since the Klein–Gordon equation is second order in the spatial derivatives, we would need two boundary conditions; namely, the continuity of the wave function and its first spatial derivative at $z = 0$:

$$\psi_1(0) = \psi_2(0) \qquad \Longrightarrow \qquad N_I + N_R = N_T, \tag{12.93}$$

$$\frac{d\psi_1(0)}{dz} = \frac{d\psi_2(0)}{dz} \qquad \Longrightarrow \qquad N_I - N_R = \frac{p_2}{p_1}N_T, \tag{12.94}$$

We need to solve for N_R and N_T such that these above two conditions are satisfied jointly. First, inserting (12.93) into (12.94), we obtain

$$N_I - N_R = \frac{p_2}{p_1}(N_I + N_R) \qquad \Longrightarrow \qquad \frac{N_R}{N_I} = \frac{p_1 - p_2}{p_1 + p_2}. \tag{12.95}$$

Second, a substitution of $N_R = N_T - N_I$ into (12.94) leads to

$$N_I - (N_T - N_I) = \frac{p_2}{p_1} N_T \qquad \Longrightarrow \qquad \frac{N_T}{N_I} = \frac{2p_1}{p_1 + p_2}. \tag{12.96}$$

To prepare the tools for studying the solutions of the Klein–Gordon equation, let us first calculate the incident, reflected, and transmitted current densities and probability densities.

Incident, reflected, and transmitted current densities
Using (12.90), we can obtain easily the incident, reflected, and transmitted current densities:

$$j_I = \frac{i\hbar}{2m}\left(\psi_I(z)\frac{d\psi_I^*(z)}{dz} - \psi_I^*(x)\frac{d\psi_I(z)}{dz}\right) = \frac{p_1}{m}|N_I|^2, \tag{12.97}$$

$$j_R = \frac{i\hbar}{2m}\left(\psi_R(z)\frac{d\psi_R^*(z)}{dz} - \psi_R^*(x)\frac{d\psi_R(z)}{dz}\right) = -\frac{p_1}{m}|N_R|^2, \tag{12.98}$$

$$j_T = \frac{i\hbar}{2m}\left(\psi_T(z)\frac{d\psi_T^*(z)}{dz} - \psi_T^*(x)\frac{d\psi_T(z)}{dz}\right) = \frac{p_2}{m}|N_T|^2,. \tag{12.99}$$

A combination of (12.97)–(12.99) yields the current densities in Regions 1 and 2

$$j_1 = j_I + j_R = \frac{p_1}{m}\left(|N_I|^2 - |N_R|^2\right), \tag{12.100}$$

$$j_2 = j_T = \frac{p_2}{m}|N_T|^2 \tag{12.101}$$

We now can obtain the reflected and transmitted coefficients by inserting (12.95) and (12.96) into (12.97)–(12.99):

$$R = \left|\frac{j_R}{j_I}\right| = \frac{|N_R|^2}{|N_I|^2} = \left|\frac{p_1 - p_2}{p_1 + p_2}\right|^2 = \frac{(p_1 - p_2)^2}{(p_1 + p_2)^2}, \tag{12.102}$$

$$T = \left|\frac{j_T}{j_I}\right| = \left|\frac{2p_2}{p_1}\right| \frac{|N_T|^2}{|N_I|^2} = \frac{4p_1p_2}{(p_1 + p_2)^2}. \tag{12.103}$$

Clearly, R and T obey the continuity condition

$$R + T = \frac{(p_1 - p_2)^2}{(p_1 + p_2)^2} + \frac{4p_1p_2}{(p_1 + p_2)^2} = 1. \tag{12.104}$$

Probability densities
Since $\Psi_I(z,\ t) = e^{-iEt/\hbar}\psi_I(z) = N_I e^{-i(Et - p_1 z)/\hbar}$, $\Psi_R(z,\ t) = e^{-iEt/\hbar}\psi_R(z) = N_R e^{-i(Et + p_1 z)/\hbar}$, and $\Psi_T(z,\ t) = e^{-i(E - V_0)t/\hbar}\psi_T(z) = N_T e^{-i[(E - V_0))t - p_2 z]/\hbar}$, we can obtain the incident, reflected, and transmitted probability densities:

$$\rho_I = \frac{i\hbar}{2mc^2}\left(\Psi_I^*\frac{\partial\Psi_I}{\partial t} - \Psi_I\frac{\partial\Psi_I^*}{\partial t}\right) = \frac{E}{mc^2}\left|\psi_I\right|^2 = \frac{E}{mc^2}\left|N_I\right|^2, \tag{12.105}$$

$$\rho_R = \frac{i\hbar}{2mc^2}\left(\Psi_R^*\frac{\partial\Psi_R}{\partial t} - \Psi_R\frac{\partial\Psi_R^*}{\partial t}\right) = \frac{E}{mc^2}\left|\psi_R\right|^2 = \frac{E}{mc^2}\left|N_R\right|^2, \tag{12.106}$$

$$\rho_T = \frac{i\hbar}{2mc^2}\left(\Psi_T^*\frac{\partial\Psi_T}{\partial t} - \Psi_T\frac{\partial\Psi_T^*}{\partial t}\right) = \frac{(E - V_0)}{mc^2}\left|\psi_T\right|^2 = \frac{(E - V_0)}{mc^2}\left|N_T\right|^2. \tag{12.107}$$

Case $V_0 - mc^2 < E < V_0 + mc^2$: Total reflection:
When $V_0 - mc^2 < E < V_0 + mc^2$ (i.e., $-mc^2 < E_2 < mc^2$), the situation becomes different from above in that the momentum in Region 2 of Figure 12.1 becomes *pure imaginary*:

$$c^2p_2^2 = E_2^2 - m^2c^4 = (E - V_0)^2 - m^2c^4 < 0 \quad \Longrightarrow \quad p_2 = ip_2', \tag{12.108}$$

where p_2' is real and positive. Hence, *the transmitted wave function* becomes an *exponentially decaying function*:

$$\psi_T(z) = N_T e^{-p_2'z/\hbar}. \tag{12.109}$$

Clearly, when p_2 is pure imaginary, the transmitted flux is zero

$$\begin{aligned} J_T &= \frac{i\hbar}{2m}\left(\psi_I(z)\frac{d}{dz}\psi_T^*(z) - \psi_I^*(z)\frac{d)}{dz}\psi_T^*(z)\right) = \frac{i\hbar}{2m}|N_T|^2\left(e^{ip_2z/\hbar}\frac{d}{dz}e^{-ip_2^*z/\hbar} - e^{-ip_2^*z/\hbar}\frac{d}{dz}e^{ip_2z/\hbar}\right) \\ &= \frac{p_2^* + p_2}{2m}|N_T|^2e^{i(p_2 - p_2^*)z/\hbar} = \frac{-ip_2' + ip_2'}{2m}|N_T|^2e^{-2p_2'z/\hbar} \\ &= 0, \end{aligned} \tag{12.110}$$

the transmitted coefficient is zero:

$$T = \left|\frac{j_T}{j_I}\right| = 0, \tag{12.111}$$

and the reflection coefficient is equal to 1. Hence, the magnitudes of the incident and reflected current densities are equal:

$$R = \left|\frac{j_R}{j_I}\right| = \left|\frac{p_1 - p_2}{p_1 + p_2}\right|^2 = \left(\frac{p_1 - p_2^*}{p_1 + p_2^*}\right)\left(\frac{p_1 - p_2}{p_1 + p_2}\right) = \frac{p_1^2 + p_2p_2^*}{p_1^2 + p_2^*p_2} = 1. \quad \Longrightarrow \quad |j_R| = |j_I|. \tag{12.112}$$

Moreover, by inserting (12.109) into (12.107), we see that, as the particle tunnels through the $z = 0$ barrier into the $z > 0$ region, the probability density decays exponentially:

$$\rho_T = \frac{(E - V_0)}{mc^2}\left|\psi_T\right|^2 = \frac{(E - V_0)}{mc^2}|N_T|^2e^{-2p_2'z/\hbar}. \tag{12.113}$$

This shows that, depending on whether E is smaller or larger than V_0, the transmitted probability density can be positive or negative:

$$\text{If} \quad V_0 < E < V_0 + mc^2 \quad \Longrightarrow \quad \rho_T > 0, \tag{12.114}$$

$$\text{If} \quad V_0 - mc^2 < E < V_0 \quad \Longrightarrow \quad \rho_T < 0. \tag{12.115}$$

The result of a negative probability density was not expected at this level. However, and as we will see below, when the potential is sufficiently strong (i.e., when $E < V_0$), a particle-antiparticle pair is created at the barrier due to the "violent" interaction of the incident flux with a strong potential that changes abruptly over a very short distance.

Case $E < V_0 - mc^2$: Klein Paradox
If the barrier is too strong, $V_0 > E + mc^2$, the momentum p_2 becomes real (see Eq. (12.91)):

$$p_2 = \pm\sqrt{\frac{E_2^2}{c^2} - m^2c^2} = \pm\sqrt{\frac{(E - V_0)^2}{c^2} - m^2c^2} \quad \Longrightarrow \quad p_2^2 > p_1^2; \tag{12.116}$$

the transmitted current density (12.99) is also real

$$j_T = \frac{p_2}{m}|N_T|^2. \tag{12.117}$$

Now, we need to make a choice on the sign of p_2. This is going to require some investigation. First, since $V_0 > E + mc^2$, the transmitted probability density (12.107) is going to be negative:

$$\rho_T = \frac{(E - V_0)}{mc^2}\left|\psi_T\right|^2 < 0. \tag{12.118}$$

Depending on the sign of p_2, the reflection coefficient (12.102) is going to be either smaller or larger than 1:

$$R = \frac{(p_1 - p_2)^2}{(p_1 + p_2)^2} \quad \Longrightarrow \quad \begin{cases} R < 1 & \text{If} \quad p_2 > 0 \\ R > 1 & \text{If} \quad p_2 < 0. \end{cases} \tag{12.119}$$

To gain more guidance on what is entering or leaving the barrier, let us calculate the group velocity v_g in Region 2:

$$(E - V_0)^2 = c^2p_2^2 + m^2c^4 \quad \Longrightarrow \quad v_g = \frac{\partial E}{\partial p_2} = \frac{p_2}{E - V_0} \quad \Longrightarrow \quad \begin{cases} v_g > 0 & \text{If } p_2 < 0 \\ v_g < 0 & \text{If } p_2 > 0. \end{cases} \tag{12.120}$$

So, for the particle to be leaving the barrier and moving to the right within Region 2 (i.e., $v_g > 0$), we must have a negative momentum, $p_2 < 0$. In this case, we have a negative current density, $j_T < 0$, which is in conformity with a negative probability density, $\rho_T < 0$:

$$p_2 < 0 \quad \Longrightarrow \quad R > 1, \quad v_g > 0, \quad \rho_T < 0, \quad j_T < 0. \tag{12.121}$$

Additionally, since the reflected current density is negative, $j_R < 0$, and bigger in magnitude than the incident current density (i.e., $|j_R| > j_I$), the continuity condition of the current density is satisfied:

$$j_I + j_R = j_T < 0. \tag{12.122}$$

In summary, the results obtained for a strong potential, $V_0 > E + mc^2$, indicate that we have a *reflection coefficient that is larger than 1* and a *reflected current density that is bigger in magnitude than the incident current density*! That is, we have *more reflected flux than the incident flux*! This *anomaly* is known as the *Klein Paradox*[22]. As we will see in Section 12.4.12, the only way to explain the $R > 1$, $\rho_T < 0$, and $j_T < 0$ anomalies is by considering the creation of particle-antiparticle pairs at the barrier where the particles propagate to the left of the barrier ($z < 0$) and the antiparticles to the right ($z > 0$), because $v_g > 0$ as shown in (12.120); while traveling to the left, the created particles increase the reflected current and thereby making it greater than the incident one. We will deal with the interpretation of these anomalies in greater details in Section 12.4.12. However, we need to mention at this level that the problems we have encountered above are due to one essential flaw: that of having treated the Klein-Gordon equation as a *single particle* equation. We have started with a *single-particle theory* – the Klein–Gordon equation – to describe the scattering of a particle from a step potential and found ourselves resorting to invoking the creation of particle-antiparticle pairs to make sense of the anomalies encountered.

[22] O. Klein, Z. *Physik*, 53, 157 (1929).

12.3.6 Klein–Gordon Equation in an Electromagnetic Field

Consider a relativistic, spinless particle of charge q which is moving under the influence of an electromagnetic field that is specified by the 4–vector potential $A^\mu = (U/c, \vec{A})$, where $U(\vec{r})$ is the electric scalar potential and $\vec{A}(\vec{r})$ is the vector potential. In this subsection, we want to derive the Klein–Gordon equation for such a particle. For this, let us recall a principle that has proven its success in non-relativistic quantum mechanics when dealing with a charged particle that interacts with an electromagnetic field. This principle states that the Hamiltonian (or Schrödinger equation) of a charged particle, q, that is interacting with an external electromagnetic field can be obtained by making the following substitutions

$$p_0^\mu \longrightarrow p^\mu = p_0^\mu - qA^\mu \qquad \text{or} \qquad \hat{H}_0 \longrightarrow \hat{H} = \hat{H}_0 - qU, \qquad \hat{\vec{p}}_0 \longrightarrow \hat{\vec{p}} = \hat{\vec{p}}_0 - q\vec{A}, \tag{12.123}$$

into the Hamiltonian $\hat{H}_0$ and momentum $\hat{\vec{p}}_0$ of the corresponding *free* particle. These expressions are known as the *minimal substitution* or the *minimal coupling* principle. In Chapter 13, we show how to obtain these relations from the classical Lagrangian of a charged particle in an electromagnetic field (see Problem 13.1). It should be noted that the minimal substitution principle has an important property: it preserves gauge invariance; it leaves the gauge symmetry of the equations intact. This is a key symmetry of the electromagnetic field.

Is the use of the minimal substitution principle justified in relativistic quantum mechanics? Since this procedure has proven its success in nonrelativistic quantum mechanics, there is no reason why it can't work in relativistic quantum theory; but this hardly constitutes a proof for its reliability. However, notwithstanding the lack of a mathematical justification for it, the reliability of this procedure is supported by extensive experimental data related to numerous nonrelativistic as well as relativistic quantum phenomena.

Note that we can express the relations (12.123) in operator form as follows

$$i\hbar\frac{\partial}{\partial t} \longrightarrow i\hbar\frac{\partial}{\partial t} - qU, \qquad -i\hbar\vec{\nabla} \longrightarrow -i\hbar\vec{\nabla} - q\vec{A}, \tag{12.124}$$

or as

$$i\hbar\partial^\mu \quad \longrightarrow \quad i\hbar\partial^\mu - qA^\mu, \tag{12.125}$$

where we have used the fact that $\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\partial_0, -\vec{\nabla}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \quad -\vec{\nabla}\right)$ and $A^\mu = (U/c, \vec{A})$.

Before embarking on dealing with the Klein–Gordon equation in an electromagnetic field, let us offer several observations about the dimensional consistency of the quantities involved in Equations (12.123) and (12.125). First, since we use the SI units in all our expressions throughout this chapter, and as shown in Appendix E (see Eqs (??)–(??)), the physical dimensions of U/c, $\vec{A}$, A_μ and qA_μ are given by

$$[U/c] = [\vec{A}] \equiv [A^\mu] = \mathrm{MLT^{-1}Q^{-1}}, \qquad [qU/c] = [qA^\mu] = \mathrm{MLT^{-1}}, \tag{12.126}$$

where M = Mass, L = Length, T = Time, and Q = electric charge. Now, since $[\hbar] = ML^2T^{-1}$ and $[\partial_\mu] = \mathrm{L^{-1}}$, we can easily infer the physical dimensions of the following important quantities that will be involved in our calculations below:

$$[i\hbar\partial^\mu - qA^\mu] = \mathrm{MLT^{-1}}; \tag{12.127}$$

that is, the physical dimensions of $(i\hbar\partial^\mu - qA^\mu)$ are those of momentum.

Now back to the Klein–Gordon equation in an electromagnetic field. Substituting (12.124) into the free particle Dirac equation (12.48), we obtain

$$\boxed{\left(i\hbar\frac{\partial}{\partial t} - qU\right)^2 \psi(\vec{r},\ t) = c^2\left(-i\hbar\vec{\nabla} - q\vec{A}\right)^2 \psi(\vec{r},\ t) + m^2c^4\psi(\vec{r},\ t).} \tag{12.128}$$

This is the Klein–Gordon equation in an electromagnetic field $A^\mu = (U/c,\ \vec{A})$.

Example 12.3
Find the 4-current j^μ for a relativistic, spinless and charged particle that is moving in an electromagnetic field $A^\mu = (U/c,\ \vec{A})$.

Solution
To find the 4-current $j^\mu = \left(c\rho(x^\mu),\ \vec{j}(x^\mu)\right)$, we need only to insert the minimal substitution relations (12.124) into (12.77). This leads to the following new expressions of the probability and current densities:

$$\rho(x^\mu) = \frac{1}{2mc^2}\left[\psi^*\left(i\hbar\frac{\partial}{\partial t} - qU\right)\psi - \psi\left(i\hbar\frac{\partial}{\partial t} + qU\right)\psi^*\right] = \rho_0(x^\mu) - \frac{1}{mc^2}qU\psi^*\psi, \tag{12.129}$$

$$\vec{j}(x^\mu) = \frac{1}{2m}\left[\psi^*\left(-i\hbar\vec{\nabla} - q\vec{A}\right)\psi - \psi\left(-i\hbar\vec{\nabla} + q\vec{A}\right)\psi^*\right] = \vec{j}_0(x^\mu) - \frac{1}{m}q\vec{A}\,\psi^*\psi \tag{12.130}$$

where ρ_0 and $\vec{j}_0$ are the probability and current densities for a charge q in the *absence* of any electromagnetic field

$$\rho_0(x^\mu) = \frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right), \qquad \vec{j}_0(x^\mu) = \frac{\hbar}{2im}\left(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right). \tag{12.131}$$

From equation (12.129), it is clear that, as in the case of a free particle, the probability density is not positive definite.

Note that we can verify that the probability and current densities (12.129) and (12.130) do verify the continuity condition:

$$\frac{\partial\rho(x^\mu)}{\partial t} + \vec{\nabla}\cdot\vec{j}(x^\mu) = 0 \qquad \Longrightarrow \qquad \partial_\mu j^\mu = 0. \tag{12.132}$$

12.3.7 Non-Relativistic Limit of the KGE in an Electromagnetic Field

We want to find out if the Klein–Gordon equation in an electromagnetic field has the correct non-relativistic limit; that is, will it reduce to the Schrödinger equation in an electromagnetic field? For this, let write the wave function $\psi(\vec{r},\ t)$ as a product a factor that describes the non-relativistic part $\phi(\vec{r},\ t)$ that involves the particle's kinetic energy and interactions and another factor $e^{-imc^2t/\hbar}$ that describes oscillations due to the particle's rest-mass energy mc^2:

$$\psi(\vec{r},\ t) = \phi(\vec{r},\ t)e^{-imc^2t/\hbar}. \tag{12.133}$$

So our task now if to focus on finding an equation for $\phi(\vec{r},\ t)$ and find out out if it is similar to the Schrödinger equation in an electromagnetic field. First, let us calculate the left-hand and right-hand terms in Eq. (12.128), separately. First, since

$$\begin{aligned}\left(i\hbar\frac{\partial}{\partial t}-qU\right)\psi(\vec{r},\ t) &= \left(i\hbar\frac{\partial}{\partial t}-qU\right)\phi(\vec{r},\ t)e^{-imc^2t/\hbar}\\ &= \left(i\hbar\frac{\partial\phi}{\partial t}-qU\phi+mc^2\phi\right)e^{-imc^2t/\hbar},\end{aligned}\tag{12.134}$$

we obtain

$$\begin{aligned}\left(i\hbar\frac{\partial}{\partial t}-qU\right)^2\psi(\vec{r},\ t) &= \left(i\hbar\frac{\partial}{\partial t}-qU\right)\left[\left(i\hbar\frac{\partial\phi}{\partial t}-qU\phi+mc^2\phi\right)e^{-imc^2t/\hbar}\right]\\ &= \left[-\hbar^2\frac{\partial^2\phi}{\partial t^2}-i\hbar qU\frac{\partial\phi}{\partial t}-i\hbar q\left(\frac{\partial U}{\partial t}\right)\phi+i\hbar mc^2\frac{\partial\phi}{\partial t}\right.\\ &\quad \left.+mc^2\left(i\hbar\frac{\partial\phi}{\partial t}-qU\phi+mc^2\phi\right)-qU\left(i\hbar\frac{\partial\phi}{\partial t}-qU\phi+mc^2\phi\right)\right]e^{-imc^2t/\hbar}\\ &= \left[-\hbar^2\frac{\partial^2\phi}{\partial t^2}-2i\hbar qU\frac{\partial\phi}{\partial t}+2i\hbar mc^2\frac{\partial\phi}{\partial t}-i\hbar q\left(\frac{\partial U}{\partial t}\right)\phi\right.\\ &\quad \left.+q^2U^2\phi-2qUmc^2\phi+m^2c^4\phi\right]e^{-imc^2t/\hbar}.\end{aligned}\tag{12.135}$$

Second, the calculation of the right-hand term in Eq. (12.128) is quite straightforward:

$$\left(-i\hbar\vec{\nabla}-q\vec{A}\right)^2\psi(\vec{r},\ t)=\left[-\hbar^2(\nabla^2\phi)+i\hbar q(\vec{\nabla}\cdot\vec{A})\phi+i\hbar q\vec{A}\cdot(\vec{\nabla}\phi)+q^2\vec{A}^{\,2}\phi\right]e^{-imc^2t/\hbar}.\tag{12.136}$$

Now, let us apply the non-relativistic limit to (12.135). In the non-relativistic limit, the particle's rest-mass energy mc^2 is much greater than all energy terms. For instance, in the energy expression $E=E_{NR}+mc^2$ where E_{NR} is the nonrelativistic part of the particle's energy (this can be the particle's kinetic energy and other energy terms, if any), we have $E_{NR}\ll mc^2$:

$$\left|\frac{\partial\phi(\vec{r},\ t)}{\partial t}\right|\sim E_{NR}\phi(\vec{r},\ t)\ll mc^2\phi(\vec{r},\ t).\tag{12.137}$$

Similarly, the following terms in (12.135) are too small than mc^2:

$$\left|\frac{\partial^2\phi}{\partial t^2}\right|\ll m^2c^4\phi,\qquad qU\left|\frac{\partial\phi}{\partial t}\right|\ll m^2c^4\phi,\qquad q^2U^2\phi\ll m^2c^4.\tag{12.138}$$

Thus, cancelling these small terms from (12.135), we obtain

$$\left(i\hbar\frac{\partial}{\partial t}-qU\right)^2\psi(\vec{r},\ t)\simeq\left[2i\hbar mc^2\frac{\partial\phi}{\partial t}-i\hbar q\left(\frac{\partial U}{\partial t}\right)\phi-2qUmc^2\phi+m^2c^4\phi\right]e^{-imc^2t/\hbar}.\tag{12.139}$$

Finally, inserting (12.136) and (12.139) into (12.128) and cancelling the exponent $e^{-imc^2t/\hbar}$ from both sides of the equation, we obtain

$$2i\hbar mc^2\frac{\partial\phi}{\partial t}-i\hbar q\left(\frac{\partial U}{\partial t}\right)\phi-2qUmc^2\phi+m^2c^4\phi\simeq c^2\left[-\hbar^2(\nabla^2\phi)+i\hbar q(\vec{\nabla}\cdot\vec{A})\phi+i\hbar q\vec{A}\cdot(\vec{\nabla}\phi)+q^2\vec{A}^{\,2}\phi\right]+m^2c^4\phi,\tag{12.140}$$

which, after carrying out few obvious simplifications, leads to

$$i\hbar\frac{\partial\phi}{\partial t} \simeq -\frac{\hbar^2}{2m}(\nabla^2\phi) + qU\phi + \frac{i\hbar q}{2m}\vec{A}\cdot(\vec{\nabla}\phi) + \frac{q^2}{2m}\vec{A}^2\phi + \frac{i\hbar q}{2m}\left[\vec{\nabla}\cdot\vec{A} + \frac{1}{c^2}\left(\frac{\partial U}{\partial t}\right)\right]\phi. \tag{12.141}$$

Now, invoking the *Lorentz gauge* condition[23]:

$$\vec{\nabla}\cdot\vec{A} + \frac{1}{c^2}\left(\frac{\partial U}{\partial t}\right) = 0, \tag{12.142}$$

we can reduce (12.141) to

$$\boxed{i\hbar\frac{\partial\phi}{\partial t} \simeq -\frac{\hbar^2}{2m}(\nabla^2\phi) + qU\phi + \frac{i\hbar q}{2m}\vec{A}\cdot(\vec{\nabla}\phi) + \frac{q^2}{2m}\vec{A}^2\phi.} \tag{12.143}$$

This is the Schrödinger equation that we have derived in Chapter 6 (see Eq. (6.196)) for a spinless particle of charge q interacting with an electromagnetic field $A^\mu = (U/c, \vec{A})$. In addition to the result we have derived in Subsection 12.3.3, this is another proof that the nonrelativisitc limit of the Klein–Gordon equation is the Schrödinger equation.

Remark:
Now that we have finished dealing with the Klein–Gordon equation, we are going to focus in the rest of this chapter on the Dirac equation. We will examine how Dirac has managed to avoid most of the problems inherent in the Klein–Gordon equation by proposing an equation which is first-order both in time and space. Although the Dirac equation yields probability densities that are positive definite, it also encounters negative energy solutions. As we will see in Subsections 12.4.10–12.4.12, much like the Klein–Gordon equation, the negative energy solutions can be properly explained only by going beyond the single particle interpretation of the Dirac equation. Dirac managed to solve the negative-energies problem by proposing the hole theory which led him to predict the existence of antiparticles; i.e., antimatter.

12.4 Dirac Equation

As discussed above, the Klein–Gordon equation suffers from two unsurmountable flaws: the occurrence of negative energy solutions and negative probability densities. These internal inconsistencies stem from its dependence on a *second order time derivative*. To avoid these problems, Dirac looked for a relativistic equation that contains only a *first order time derivative*.

Like Schroödinger who started from a relation with first order time derivatives

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = \hat{H}_S|\psi\rangle, \tag{12.144}$$

where $\hat{H}_S = -\frac{\hbar^2}{2m}\nabla^2$ for a nonrelativistic free particle (or $\hat{H}_S = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}$ for a particle moving in a potential $\hat{V}$), and unlike Klein and Gordon who started from a relation with second order time derivatives

$$-\hbar^2\frac{\partial^2}{\partial t^2}|\psi\rangle = \hat{H}_{KG}^2|\psi\rangle, \tag{12.145}$$

[23] As a reminder, we are using the SI units in all equations throughout this chapter. Moreover, it helps to recall that the Lorentz gauge condition is given in the Gaussian (cgs) units by: $\vec{\nabla}\cdot\vec{A} + \frac{1}{c}\left(\frac{\partial U}{\partial t}\right) = 0$.

where $\hat{H}_{KG}^2 = -\hbar^2c^2\nabla^2 + m^2c^4$ for a relativistic free particle, Dirac started with a wave equation that involves only first order time derivatives[24]:

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = \hat{H}_D|\psi\rangle, \tag{12.146}$$

where the Dirac Hamiltonian $\hat{H}_D$ has yet to be specified. In addition to making his equation first-order in time, Dirac imposed a second requirement: the solutions of his equation must satisfy the Klein–Gordon equation; as will be seen below, this requirement will play a crucial role in obtaining the proper expression of $\hat{H}_D$.

It helps to note that, since the Dirac equation accounts for the spin, the state function $|\psi\rangle$ of (12.146) cannot be a scalar; it is expected to have a number of components, very much like the spinor states that we have encountered in Chapter 6 (see Eq. (6.4) on Page 403 when we have included the spin in the Shcrödinger's nonrelativistic equation.

12.4.1 Dirac Equation for a Free Particle

For the sake of simplicity, we are going to illustrate the most essential elements of the Dirac equation just by considering an isolated free particle. As in non-relativistic quantum mechanics, the free particle provides a simple, yet highly edifying, platform within which one can illustrate a number of key aspects of the theory without undue mathematical complications. One can learn a great deal from the free particle example so as to inform our approach when one looks at the solutions of the Dirac equation for more involved examples such as particles under the influence of potentials.

Unlike the asymmetrical dependence of the Shrödinger equation on time and space (first order in time derivatives and second order in spatial derivatives), and unlike the symmetrical dependence on time and space (second order both in time derivatives and spatial derivatives) of the Klein–Gordon equation, *Dirac looked for a wave equation with a symmetrical dependence both on time and space (first order both in time derivatives and spatial derivatives) so as to be proportional to* $\partial_\mu = \left((1/c)\partial/\partial t,\ \vec{\nabla}\right)$ and, hence, *covariant; special relativity requires treating space and time on equal footing*. Guided by this requirement, Dirac posited a Hamiltonian which is first order in spatial derivatives,

$$\hat{H} = c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2, \tag{12.147}$$

and wrote his equation, the Dirac equation for a free particle, in the following form:

$$\hat{H}|\psi\rangle = \left(c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2\right)|\psi\rangle, \tag{12.148}$$

where the coefficients $\vec{\alpha}$ and β are to be determined; we will be denoting the components of $\vec{\alpha}$ by $\vec{\alpha} = (\alpha_x,\ \alpha_y,\ \alpha_z) \equiv (\alpha_1,\ \alpha_2,\ \alpha_3)$ or α_j $(j = 1,\ 2,\ 3)$. The presence of the factors c and c^2 in (12.148) is intended to make the coefficients α_j and β dimensionless; the physical dimensions of $c\vec{p}$ and mc^2 in (12.148) are those of energy as it is the case for $\hat{H}$ on the left hand side. Additionally, since all points in spacetime are equivalent for free particles (the Hamiltonian

[24]Note that, unlike Schrödinger and Klein–Gordon equations which are both second order in the spatial derivatives, the Dirac equation involves only first order derivatives in the spatial coordinates; by treating time and the spatial coordinates symmetrically (i.e., involving only first order derivatives both in time and the spatial coordinates), Dirac ensured that his equation is Lorentz invariant.

is invariant under spacetime displacements), the coefficients α_j and β must be constant; they must be independent of the variables $\vec{r}$, $\vec{p}$ and time, t. Although there is no need to require α_j and β to commute with each other (as will be seen below), they must commute with $\vec{r}$ and $\vec{p}$. To ensure that the Hamiltonian is hermitian, the constants α_j and β must also be hermitian:

$$\alpha_j^\dagger = \alpha_j, \qquad \beta^\dagger = \beta. \tag{12.149}$$

For notational brevity, we may sometimes denote x^μ and $\psi(x^\mu)$ by x and $\psi(x)$, respectively. We can rewrite (12.148) in the position representation as follows:

$$\boxed{i\hbar\frac{\partial}{\partial t}\psi(x^\mu) = \left(-i\hbar c\vec{\alpha}\cdot\vec{\nabla} + mc^2\beta\right)\psi(x^\mu).} \tag{12.150}$$

The Dirac equation is sometimes described as the square root of the Klein–Gordon equation because squaring the Dirac equation results in the Klein–Gordon equation. Hence, any state $|\psi\rangle$ that solves the Dirac equation must also obey the Klein–Gordon equation. This is the key requirement we are going to use to find the constants α_j and β; namely, we can determine them by demanding that $\psi(x^\mu)$ of (12.150) must also satisfy the Klein–Gordon equation for a free particle. This reduces to the requirement that, upon squaring the Dirac Hamiltonian $\hat{H} = c\vec{\alpha}\cdot\hat{\vec{p}} + mc^2\beta$, the Dirac equation must give the correct energy–momentum expression, $E^2 = c^2\vec{p}^{\,2} + m^2c^4$. The square of (12.150),

$$-\hbar^2\frac{\partial^2}{\partial t^2}\psi(x^\mu) = \left(c\vec{\alpha}\cdot\hat{\vec{p}} + mc^2\beta\right)^2\psi(x^\mu), \tag{12.151}$$

does lead to an equation with a structure that is similar to the Klein–Gordon equation (12.49), where

$$\begin{aligned}\left(c\vec{\alpha}\cdot\hat{\vec{p}} + mc^2\beta\right)^2 &= c^2\left(\alpha_x\hat{p}_x + \alpha_y\hat{p}_y + \alpha_z\hat{p}_z + \beta mc\right)^2 \\ &= c^2\left(\alpha_x^2\hat{p}_x^2 + \alpha_y^2\hat{p}_y^2 + \alpha_z^2\hat{p}_z^2\right) + \beta^2m^2c^4 \\ &\quad + c^2\Big[(\alpha_x\alpha_y + \alpha_y\alpha_x)\hat{p}_x\hat{p}_y + (\alpha_x\alpha_z + \alpha_z\alpha_x)\hat{p}_x\hat{p}_z + (\alpha_y\alpha_z + \alpha_z\alpha_y)\hat{p}_y\hat{p}_z\Big] \\ &\quad + mc^3\Big[(\alpha_x\beta + \beta\alpha_x)\hat{p}_x + (\alpha_y\beta + \beta\alpha_y)\hat{p}_y + (\alpha_z\beta + \beta\alpha_z)\hat{p}_z\Big].\end{aligned} \tag{12.152}$$

Now, by equating (12.152) to $E^2 = c^2\vec{p}^{\,2} + m^2c^4$, we see that the square of the constants α_j and β is unity and that they anticommute in pairs

$$\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = 1, \qquad \beta^2 = 1, \tag{12.153}$$

$$\alpha_i\alpha_j + \alpha_j\alpha_i = \left\{\alpha_i, \alpha_j\right\} = 2\delta_{ij}, \qquad i, j = 1, 2, 3, \tag{12.154}$$

$$\alpha_j\beta + \beta\alpha_j = \left\{\alpha_j, \beta\right\} = 0, \qquad j = 1, 2, 3\,; \tag{12.155}$$

the anticommutator between any two operators (or matrices) $\hat{A}$ and $\hat{B}$ is defined by: $\left\{\hat{A}, \hat{B}\right\} = \hat{A}\hat{B} + \hat{B}\hat{A}$. Since the components α_j ($j = 1, 2, 3$) do not commute with each other as well with β

(they actually anticommute), they cannot be numbers, they must be *matrices*; their sizes have yet to be determined. However, since $\alpha_j^2 = \beta^2 = 1$, the coefficients α_j and β must be *square* matrices with eigenvalues ± 1; they are known as the Dirac matrices. It is also worth noticing that the traces of α_j and β are equal to zero due to their anticommution properties listed in equations (12.154) and (12.155):

$$\mathrm{Tr}(\beta) = \mathrm{Tr}(\alpha_j) = 0, \qquad j = 1,\ 2,\ 3. \tag{12.156}$$

The set of algebraic relations (12.154) – (12.155) are known as the *Dirac algebra* which, in turn, is a special case of the *Clifford algebra*.

12.4.2 Representation of the Dirac Matrices

Let us now focus on finding a representation for the Dirac matrices that will satisfy the anticommutation relations (12.154) – (12.155). Since α_j and β are square matrices, say of $N \times N$ dimension, and since $\alpha_i\alpha_j = -\alpha_j\alpha_i$ for $i \neq j$ (see (12.154)), we can show that[25]

$$\det(\alpha_i\alpha_j) = \det(-\alpha_j\alpha_i) = (-1)^N \det(\alpha_i\alpha_j) \qquad \Longrightarrow \qquad (-1)^N = 1, \tag{12.157}$$

and, hence, N must be *even*. The eveness of N can also be seen as follows: as discussed above, since α_i and β are square matrices, traceless, and have eigenvalues ± 1, they must have even dimensions.

Notice that, since α_j and β are $N \times N$ matrices, the wave function that is solution to Dirac's equation (12.150) must be a *column vector or spinor* with an *even* number of N components[26]

$$\psi(x^\mu) = \langle x^\mu|\psi\rangle = \begin{pmatrix} \psi_1(x^\mu) \\ \psi_2(x^\mu) \\ \vdots \\ \psi_N(x^\mu) \end{pmatrix}; \tag{12.158}$$

the number N has yet to be found. Since N must be even, starting with the lowest value, $N = 2$, a suitable representation of the Dirac matrices can be provided by Pauli's three matrices σ_x, σ_y, σ_z, for they are hermitian, mutually anticommuting, and they satisfy the needed conditions (12.153) – (12.155):

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I_2, \qquad \{\sigma_x,\ \sigma_y\} = \{\sigma_x,\ \sigma_z\} = \{\sigma_y,\ \sigma_z\} = 0, \tag{12.159}$$

where I_2 is the 2×2 unity matrix. As seen in Chapter 5, the above relations can be summarized in the following anticommutation relation

$$\{\sigma_j, \sigma_k\} = 2I_2\delta_{j,k}, \qquad (j,\ k = x,\ y,\ z), \tag{12.160}$$

which defines the *three-dimensional Clifford algebra* of the Pauli matrices. But we need four matrices, not three; it is simply not possible to find a fourth 2×2 matrix that anticommutes

[25] We have made use of these two theorems from linear algebra: (1) If A is an $N \times N$ matrix and k a constant, then $\det(kA) = k^N \det(A)$, (2) If A and B are two square matrices, then: $\det(AB) = \det(A)\det(B) = \det(B)\det(A) = \det(BA)$.

[26] Note that each one of the N components of $\psi(x^\mu)$ must satisfy the Klein–Gordon equation so as to fulfill the energy–momentum relation $E^2 = p^2c^2 + m^2c^4$.

with all the matrices σ_x, σ_y, σ_z. Hence, the $N = 2$ possibility is ruled out and that the lowest value of N must then be 4 (i.e., $N \geq 4$). Therefore, the α_j and β must be at least 4×4 matrices. This "coincidence" is quite interesting: since the smallest representation of α_j and β is four-dimensional, we might consider the three dimensions for space and one for time! As we see shortly (see (12.181)), the smallest non-trivial representation of the Clifford algebra for the α_j and β matrices is four-dimensional. Obviously, the representation of the Clifford algebra strongly depends on the dimension and metric of spacetime.

In sum, we are looking for a representation of the α_j and β matrices that must be hermitian, square, mutually anticommuting, traceless, that have ± 1 eigenvalues, with a minimum dimension of 4×4, and which must satisfy the algebraic relations (12.154) – (12.155). One can find numerous sets of 4×4 matrices that satisfy these requirements; all these choices will lead to the same physical outcomes. Hence, any such set will be satisfactory. Using the three Pauli matrices plus a fourth one for β, Dirac has constructed a representation of the α_j and β matrices as follows:

$$\beta = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}, \qquad \vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \tag{12.161}$$

where I_2 and $\vec{\sigma}$ are the 2×2 unit and Pauli matrices, respectively:

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{12.162}$$

Combining (12.161) and (12.162), we can infer the 4×4 matrices for β and α_j:

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \alpha_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \alpha_y = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \alpha_z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \tag{12.163}$$

The four-dimensional representation (12.161) – (12.163), called the Dirac representation, is suitable for spin 1/2 particles. For particles with spins larger than 1/2, one needs to go to higher representations where $N \geq 6$.

So, in the four-dimensional representation (12.161) – (12.163, the state $|\psi\rangle$ is represented by a four-component column matrix, called a 4-component spinor:

$$\psi(x^\mu) = \begin{pmatrix} \psi_1(x^\mu) \\ \psi_2(x^\mu) \\ \psi_3(x^\mu) \\ \psi_4(x^\mu) \end{pmatrix}, \tag{12.164}$$

where $\psi_j(x^\mu)$ $(j = 1, 2, 3, 4)$ are complex functions. It is important to keep in mind that this 4-component spinor is not a 4-vector; the indexes of the four components ψ_1, ψ_2, ψ_3, ψ_4 have nothing to do with the x, y, z, t spacetime components of 4-vectors.

Now, we can write the Dirac equation in matrix form in two, equivalent, ways. First, by inserting (12.161) into (12.150), we have:

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = \begin{pmatrix} mc^2 I_2 & -i\hbar c \vec{\sigma} \cdot \vec{\nabla} \\ -i\hbar c \vec{\sigma} \cdot \vec{\nabla} & -mc^2 I_2 \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix}, \qquad \text{with} \qquad \phi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \quad \chi = \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix}, \tag{12.165}$$

where ϕ and χ are two-component spinors. Second, by inserting (12.164) into (12.150), we obtain four Dirac equations, one for each component $\psi_j(x)$ where $j = 1, 2, 3, 4$:

$$i\hbar\frac{\partial}{\partial t}\psi_j(x^\mu) = \left[-i\hbar c\left(\alpha_x\frac{\partial}{\partial x} + \alpha_y\frac{\partial}{\partial y} + \alpha_z\frac{\partial}{\partial z}\right) + mc^2\beta\right]\psi_j(x^\mu); \tag{12.166}$$

so, the Dirac equation consists of four coupled differential equations which can be written in matrix form as follows

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix} = \begin{pmatrix} mc^2 & 0 & -i\hbar c\frac{\partial}{\partial z} & -\hbar c\left(i\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)\\ 0 & mc^2 & -\hbar c\left(i\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right) & i\hbar c\frac{\partial}{\partial z}\\ -i\hbar c\frac{\partial}{\partial z} & -\hbar c\left(i\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) & -mc^2 & 0\\ -\hbar c\left(i\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right) & i\hbar c\frac{\partial}{\partial z} & 0 & -mc^2\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix}. \tag{12.167}$$

It helps to notice that each components ψ_j satisfies the Klein–Gordon equation. We also should note that, while the Dirac spinor consists of a four-column matrix, the Dirac Hamiltonian is represented by a 4×4 matrix.

Notice that we can write the Dirac Hamiltonian (12.167) in terms of the momentum operators $\hat{p}_x$, $\hat{p}_y$, $\hat{p}_z$ as follows

$$\hat{H} = \begin{pmatrix} mc^2 & 0 & c\hat{p}_z & c(\hat{p}_x - i\hat{p}_y)\\ 0 & mc^2 & c(\hat{p}_x + i\hat{p}_y) & -c\hat{p}_z\\ c\hat{p}_z & c(\hat{p}_x - i\hat{p}_y) & -mc^2 & 0\\ c(\hat{p}_x + i\hat{p}_y) & -c\hat{p}_z & 0 & -mc^2\end{pmatrix}. \tag{12.168}$$

It is worth mentioning that, unlike the Schrödinger and the Klein–Gordon wave functions which have only one single component each, and unlike the Pauli spinors that have two components (due to the two possible spin orientations for a spin-1/2 particle), the Dirac state function $|\psi\rangle$ has four components. As we will see below, two components describe the two possible spin orientations of a spin-1/2 particle and the two other components to account for the two spin orientations of its antiparticle. The four components of the spinor are a direct consequence of the Dirac equation being first order in time and space derivatives to account for the one time component and the three spatial components. As seen above, Pauli has used two-dimensional matrices, while Dirac has used four-dimensional matrices which are expressed in terms of three Pauli matrices plus a fourth matrix which is essentially the unit matrix.

Now that we have specified the representation of the Dirac matrices and, hence, the dimensions of the Dirac spinor and Hamiltonian, we are ready to look at the solutions of the Dirac equation and their implications on important issues such as the continuity equation, charge and current conservation, positive definiteness of the probability density, and treatment of the spin. However, before delving into these issues, we need to derive the covariant form of the Dirac equation; the covariant notation will make the presentation of these issues much transparent and more elegant.

12.4.3 Dirac Matrices and their Properties

To obtain the covariant form of the Dirac equation, let us introduce a new set of matrices, known as the Dirac γ^μ matrices, that are more convenient than the α_j and β matrices introduced

above:

$$\gamma^\mu = \left(\gamma^0, \vec{\gamma}\right) = (\beta, \beta\vec{\alpha}) = \left(\gamma^0, \gamma^1, \gamma^2, \gamma^3\right), \quad \gamma^0 = \beta, \quad \gamma^j = \beta\alpha_j, \quad j = 1, 2, 3, \tag{12.169}$$

where γ^0 is t he time-like component and the three matrices γ^1, γ^2, γ^3 are the space-like components. It is important to note that the *Dirac γ^μ matrices are not 4-vectors*, they are constant matrices and, hence, they are invariant under Lorentz transformations.

The properties of the γ^μ matrices can be deduced from those of β and α_j listed in (12.149) and (12.153)–(12.155). First, since β and α_j are Hermitian, we can show that γ^0 is Hermitian, but γ^j is anti-Hermitian:

$$(\gamma^0)^\dagger = \gamma^0, \qquad (\gamma^j)^\dagger = -\gamma^j, \qquad j = 1, 2, 3. \tag{12.170}$$

The proof is quite straightforward: $(\gamma^0)^\dagger = \beta^\dagger = \beta = \gamma^0$ and $(\gamma^j)^\dagger = (\beta\alpha_j)^\dagger = \alpha_j^\dagger\beta^\dagger = \alpha_j\beta = -\beta\alpha_j = -\gamma^j$, since $\alpha_j\beta = -\beta\alpha_j$. We can rewrite (12.170) in the following simpler form

$$(\gamma^\mu)^\dagger = \gamma_\mu = \left(\gamma^0, -\gamma^1, -\gamma^2, -\gamma^3\right) = \left(\gamma^0, -\vec{\gamma}\right), \qquad \mu = 0, 1, 2, 3. \tag{12.171}$$

Second, using equations (12.153)–(12.155), we can infer:

$$\left(\gamma^0\right)^2 = I_4, \qquad \left(\gamma^j\right)^2 = -I_4, \qquad j = 1, 2, 3, \tag{12.172}$$

where I_4 is the four-rowed identity matrix. While the proof of the first equation is self-evident (i.e., $\left(\gamma^0\right)^2 = \beta^2 = I_4$), the proof of the second equation is also easy to see:

$$\left(\gamma^j\right)^2 = \left(\beta\alpha_j\right)\left(\beta\alpha_j\right) = \beta\left(\alpha_j\beta\right)\alpha_j = -\beta\left(\beta\alpha_j\right)\alpha_j = -\beta^2\alpha_j^2 = -I_4. \tag{12.173}$$

Note that, since $\left(\gamma^0\right)^2 = I_4$ and $\left(\gamma^j\right)^2 = -I_4$ (where $j = 1, 2, 3$), then γ^0 and γ^j have real and imaginary eigenvalues, respectively:

$$\left(\gamma^0\right)^2 = I_4 \implies \gamma^0 \text{ has real eigenvalues } \pm 1, \tag{12.174}$$

$$\left(\gamma^j\right)^2 = -I_4 \implies \gamma^j \text{ has imaginary eigenvalues } \pm i. \tag{12.175}$$

Additionally, we can show that (12.170) can be condensed into the following lighter form

$$(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0, \qquad \mu = 0, 1, 2, 3. \tag{12.176}$$

This relation can be verified as follows: when $\mu = 0$, we have $\gamma^0\gamma^0\gamma^0 = \beta^3 = \beta^2\beta^0 = \beta^0 = \gamma^0 \equiv \left(\gamma^0\right)^\dagger$ and when $\mu = j$ (where $j = 1, 2, 3$), we have $\gamma^0\gamma^j\gamma^0 = \beta\left(\beta\alpha_j\right)\beta = \beta\beta\left(\alpha_j\beta\right) = -\beta^3\alpha_j = -\beta\alpha_j = -\gamma^j \equiv \left(\gamma^j\right)^\dagger$. From (12.176) we can obtain

$$\gamma^\mu = \gamma^0(\gamma^\mu)^\dagger\gamma^0, \qquad \mu = 0, 1, 2, 3. \tag{12.177}$$

Now, since $\{\beta, \beta\} = 2I_4$, we have

$$\left\{\gamma^0, \gamma^0\right\} = 2I_4. \tag{12.178}$$

Similarly, since $\alpha_j\beta = -\beta\alpha_j$, $\beta^2 = I_4$, and $\{\alpha_i, \alpha_j\} = 2\delta_{ij}I_4$, we obtain

$$\begin{aligned}\{\gamma^i, \gamma^j\} &= \{\beta\alpha_i, \beta\alpha_j\} = \beta\alpha_i\beta\alpha_j + \beta\alpha_j\beta\alpha_i = -\beta^2\alpha_i\alpha_j - \beta^2\alpha_j\alpha_i \\ &= -\{\alpha_i, \alpha_j\} \\ &= -2\delta_{ij}I_4. \end{aligned} \tag{12.179}$$

Also, since $\alpha_j\beta = -\beta\alpha_j$ and $\beta^2 = I_4$, we have

$$\begin{aligned}\{\gamma^0, \gamma^j\} &= \{\beta, \beta\alpha_j\} = \beta^2\alpha_j + \beta\alpha_j\beta = \alpha_j - \beta^2\alpha_j = \alpha_j - \alpha_j \\ &= 0, \qquad j = 1, 2, 3. \end{aligned} \tag{12.180}$$

Hence, we can condense the last three equations (12.178) – (12.180) into the following relation

$$\boxed{\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4, \qquad \mu, \nu = 0, 1, 2, 3,} \tag{12.181}$$

where $g^{\mu\nu}$ is the Minkowski metric (+, −, −, −) defined in (12.18); i.e., $g^{00} = 1$ and $g^{11} = g^{22} = g^{33} = -1$ and $g^{\mu\nu} = 0$ when $\mu \neq \nu$. The anticommutator (12.181), which is nothing but a condensed version of the algebraic relations (12.154) – (12.155), defines the *four-dimensional Clifford algebra* for the γ–matrices; the elements of this four-dimensional vector space are called spinors or Dirac 4-spinors. It is worth mentioning that, due to their anticommutation property (12.181), the γ^μ matrices are traceless, very much like the β and α_j matrices (see (12.156)).

Up to this point, we have defined the γ–matrices with upper indices only, γ^μ. The Minkowski metric allows us to define the lower index matrix γ_μ as follows

$$\gamma_\mu = g_{\mu\nu}\gamma^\nu, \qquad \mu, \nu = 0, 1, 2, 3. \tag{12.182}$$

As shown in Problem 12.14, the "covariant" γ–matrices, γ_0, γ_1, γ_2, γ_3, are related to their "contravariant" counterparts, γ^0, γ^1, γ^2, γ^3, by

$$\gamma_0 = \gamma^0, \qquad \gamma_1 = -\gamma^1, \qquad \gamma_2 = -\gamma^2, \qquad \gamma_3 = -\gamma^3. \tag{12.183}$$

Multiplying both sides of (12.181) by $g_{\mu\nu}$, we obtain

$$g_{\mu\nu}(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu) = 2g_{\mu\nu}g^{\mu\nu} \quad \Longrightarrow \quad \gamma^\mu\gamma_\mu = 4I_4, \tag{12.184}$$

where we have used the fact that $g_{\mu\nu}g^{\mu\nu} = 4$ and that $g_{\mu\nu}\gamma^\mu\gamma^\nu = \gamma^\mu\gamma_\mu$ and $g_{\mu\nu}\gamma^\nu\gamma^\mu = \gamma^\mu\gamma_\mu$; another proof is provided in Example 12.4. Additionally, we can show the following useful relations:

$$\gamma^\nu\gamma^\mu\gamma_\nu = -2\gamma^\mu, \qquad \gamma^\nu\gamma^\mu\gamma^\lambda\gamma_\nu = 4g^{\mu\lambda}I_4, \qquad \gamma^\nu\gamma^\mu\gamma^\lambda\gamma^\sigma\gamma_\nu = -2\gamma^\sigma\gamma^\lambda\gamma^\mu. \tag{12.185}$$

We may introduce a useful quantity (a rank 2 antisymmetric tensor) which is constructed in terms of the γ–matrices as follows:

$$\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \quad \gamma^\nu]; \tag{12.186}$$

these are known as the generators of the proper orthochronous Lorentz transformations; we will use these transformations below (see equation (12.255)) to prove the Lorentz covariance of the Dirac equation. Note that the Hermiticity relation for $\sigma^{\mu\nu}$ is similar to that of γ^μ; namely, using (12.176), we have

$$(\sigma^{\mu\nu})^\dagger = \gamma^0\sigma^{\mu\nu}\gamma^0, \qquad \mu = 0,\ 1,\ 2,\ 3. \tag{12.187}$$

We may now introduce a fifth matrix which is defined by the product of the four γ−matrices:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3, \tag{12.188}$$

where γ^5 is not to be viewed as one of the ordinary γ matrices; the number "5" in γ^5 is neither a component nor an index, it is simply a notational label. The ordinary four matrices γ^0, γ^1, γ^2, γ^3 pertain to the four components of the gamma matrix γ^μ within the 4-dimensional space-time. Note that the Greek indices μ, ν, α, β, etc will always stand for the four values 0, 1, 2, 3 only, and not for 4 nor for 5.

We may also define γ_5 as follows

$$\gamma_5 = \frac{i}{4!}\varepsilon_{\alpha\beta\mu\nu}\gamma^\alpha\gamma^\beta\gamma^\mu\gamma^\nu \equiv -i\gamma_0\gamma_1\gamma_2\gamma_3 = \gamma^5, \tag{12.189}$$

where the antisymmetric tensor $\varepsilon_{\alpha\beta\mu\nu}$ is defined by

$$\varepsilon_{\alpha\beta\mu\nu} = \begin{cases} 1, & \text{for } \alpha,\ \beta,\ \mu,\ \nu \text{ an even permutation of } 0,\ 1,\ 2,\ 3, \\ 0, & \text{if any two or more indices among } \alpha,\ \beta,\ \mu,\ \nu \text{ are equal}, \\ -1, & \text{for } \alpha,\ \beta,\ \mu,\ \nu \text{ an odd permutation of } 0,\ 1,\ 2,\ 3. \end{cases} \tag{12.190}$$

Using (12.188) and (12.183), we can verify the relation $\gamma^5 = \gamma_5$ as follows:

$$\gamma_5 = -i\gamma_0\gamma_1\gamma_2\gamma_3 = -i(+\gamma^0)(-\gamma^1)(-\gamma^2)(-\gamma^3) \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \gamma^5. \tag{12.191}$$

We can show that

$$\left(\gamma^5\right)^2 = \gamma_5^2 = I_4, \qquad \left(\gamma^5\right)^\dagger = \gamma^5, \qquad \gamma_5^\dagger = \gamma_5; \tag{12.192}$$

hence, both of γ^5 and γ_5 have real eigenvalues ± 1. We can also show that γ^5 and γ_5 anticommute with γ^μ:

$$\left\{\gamma^5,\ \gamma^\mu\right\} = 0, \qquad \{\gamma_5,\ \gamma^\mu\} = 0, \qquad \mu = 0,\ 1,\ 2,\ 3. \tag{12.193}$$

Using (12.186), we can easily show that γ_5 commutes with $\sigma^{\mu\nu}$:

$$\begin{aligned} [\gamma_5,\ \sigma^{\mu\nu}] &= \frac{i}{2}\left(\gamma_5\gamma^\mu\,\gamma^\nu - \gamma_5\gamma^\nu\,\gamma^\mu - \gamma^\mu\gamma^\nu\gamma_5 + \gamma^\nu\gamma^\mu\gamma_5\right) \\ &= \frac{i}{2}\left(\gamma_5\gamma^\mu\,\gamma^\nu - \gamma_5\gamma^\nu\,\gamma^\mu - \gamma_5\gamma^\mu\,\gamma^\nu + \gamma_5\gamma^\nu\,\gamma^\mu\right) \\ &= 0, \end{aligned} \tag{12.194}$$

where, in the middle line, we have used the relations $\gamma^\mu\gamma^\nu\gamma_5 = -\gamma^\mu\,\gamma_5\gamma^\nu = \gamma_5\gamma^\mu\,\gamma^\nu$ and $\gamma^\nu\gamma^\mu\gamma_5 = -\gamma^\nu\gamma_5\gamma^\mu = \gamma_5\gamma^\nu\gamma^\mu$ which can be obtained from $\gamma_5\gamma^\mu = -\gamma^\mu\gamma_5$ (see Eq. (12.193)).

It is worth mentioning that, since the Dirac matrices are constants, they are the same in all inertial frames. However, when contracted with $\overline{\psi}$ and ψ, different bilinear combinations have

Table 12.2: The 16 linearly independent Γ matrices that can be constructed from the various products of the γ matrices.

Γ	Type	Number of elements
1	scalar	1
γ^μ	vector	4
γ^5	pseudo-scalar	1
$\gamma^\mu\gamma^5$	pseudo-vector	4
$\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$	tensor	6

their own distinct transformation properties; the bilinear quantities $\overline{\psi}\psi$, $\overline{\psi}\gamma^\mu\psi$, $\overline{\psi}\gamma^5\psi$, $\overline{\psi}\gamma^\mu\gamma^5\psi$, $\overline{\psi}\sigma^{\mu\nu}\psi$ are all covariant. By taking products of the various γ matrices, we can construct 16 linearly independent 4×4 matrices which we denote by Γ_n (where $n = 1, 2, \cdots, 16$) or, in short, Γ. Any product of the γ matrices can be expressed as a linear combination of the 16 matrices displayed in Table 12.2. This means that the set $\left\{1, \gamma^\mu, \gamma^5, \gamma^\mu\gamma^5, \sigma^{\mu\nu}\right\}$ forms a complete basis for any 4×4 matrix. As we will see below (see Subsection 12.4.23), the prefix "pseudo" in pseudo-scalars and pseudo-vectors refers to the way these quantities transform under parity, $\vec{r} \longrightarrow -\vec{r}$. While scalars remain unchanged under parity, pseudo-scalars change sign; however, the space components of vectors change sign under parity, whereas those of pseudo-vectors do not. We should note that all 16 matrices defined in Table 12.2 satisfy this relation: $\Gamma^2 = I_4$.

The inner product of a 4-vector b_μ by γ^μ is often encountered; it is given by

$$\begin{aligned} b_\mu\gamma^\mu &= b^\mu\gamma_\mu = b^0\gamma^0 - \vec{b}\cdot\vec{\gamma} = b^0\gamma^0 - b^1\gamma^1 - b^2\gamma^2 - b^3\gamma^3 \\ &= b^0\gamma^0 - b_x\gamma^1 - b_y\gamma^2 - b_z\gamma^3, \end{aligned} \tag{12.195}$$

where $\gamma^\mu = \left(\gamma^0, \vec{\gamma}\right) \equiv \left(\gamma^0, \gamma^1, \gamma^2, \gamma^3\right)$ and $b^\mu = \left(b^0, \vec{b}\right) \equiv \left(b^0, b^1, b^2, b^3\right)$ or $b^\mu \equiv \left(b^0, b_x, b_y, b_z\right)$; here b^0, b_x, b_y, b_z can be 4×4 matrices, much like γ^0, γ^1, γ^2, γ^3, or simply row vectors (i.e., 1×4 vectors) that get contracted with γ^μ.

It helps to introduce the *Fenyman slash* notation which provides an abbreviation of $\gamma-$matrices with 4-vectors; it is commonly encountered in the covariant formulation of the Dirac equation. This notation gives the product of a $\gamma-$matrix with a 4-vector b^μ:

$$\gamma^\mu b_\mu \equiv \not{b} = \gamma^0 b^0 - \vec{\gamma}\cdot\vec{b} = \gamma^0 b^0 - \gamma^1 b_x - \gamma^2 b_y - \gamma^3 b_z, \tag{12.196}$$

where, as shown in Problem (12.13) on Page 838, *the slashed quantity $\not{b}$ represents a 4×4 matrix*, since γ^μ is a 4×4 matrix. In particular, we can write the inner product of γ^μ with the 4-gradient operator and the 4-momentum operator as follows:

$$\gamma^\mu\partial_\mu \equiv \not{\partial} = \frac{\gamma^0}{c}\frac{\partial}{\partial t} + \vec{\gamma}\cdot\vec{\nabla}, \tag{12.197}$$

and

$$\gamma^\mu\hat{p}_\mu \equiv \not{\hat{p}} = i\hbar\not{\partial} = \frac{i\hbar\gamma^0}{c}\frac{\partial}{\partial t} + i\hbar\vec{\gamma}\cdot\vec{\nabla} \tag{12.198}$$

From the definition (12.196), we can show the following relations:

$$\gamma_\mu \not{b} \gamma^\mu = -2\not{b}, \qquad \gamma_\mu \not{a} \not{b} \gamma^\mu = 4a \cdot b \equiv 4a^\mu b_\mu, \qquad \gamma_\mu \not{a} \not{b} \not{c} \gamma^\mu = -2\not{c} \not{b} \not{a}, \tag{12.199}$$

$$\not{a}\not{a} = a^2 \equiv a_\mu a^\mu, \qquad \not{a}\not{b} = a \cdot b - i a_\mu \sigma^{\mu\nu} b_\nu \equiv a_\mu b^\mu - i a_\mu \sigma^{\mu\nu} b_\nu, \tag{12.200}$$

where a, b, c are 4-vectors and $\sigma^{\mu\nu}$ is given by (12.186).

Some useful trace theorems involving the γ^μ matrices

Using the various properties of the γ–matrices listed above (most notably equations (12.184), (12.185), (12.193), (12.199) and (12.200)), we can prove a number of trace theorems. For instance, and as discussed above, due to their anticommutation properties (see (12.181) and (12.193)), the γ^μ matrices are traceless:

$$\text{Tr}\, \gamma^\mu = 0, \qquad \text{Tr}\, \gamma^5 = 0, \qquad \mu = 0,\ 1,\ 2,\ 3. \tag{12.201}$$

Evidently, the trace of the 4-rowed identity matrix is 4:

$$\text{Tr}\, I_4 = 4. \tag{12.202}$$

The trace of the product of a number of γ matrices is invariant under their *cyclic permutations*[27]

$$\begin{aligned} \text{Tr}\,(\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{n-1}}\gamma^{\mu_n}) &= \text{Tr}\,(\gamma^{\mu_2}\cdots\gamma^{\mu_{n-1}}\gamma^{\mu_n}\gamma^{\mu_1}) = \text{Tr}\,(\gamma^{\mu_3}\cdots\gamma^{\mu_{n-1}}\gamma^{\mu_n}\gamma^{\mu_1}\gamma^{\mu_2}) \\ &= \cdots = \text{Tr}\,(\gamma^{\mu_n}\gamma^{\mu_{n-1}}\cdots\gamma^{\mu_2}\gamma^{\mu_1}). \end{aligned} \tag{12.203}$$

Another useful theorem: the trace of the product of an *odd number* of γ^μ matrices is *zero*:

$$\text{Tr}\,(\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_3}\cdots\gamma^{\mu_{2n-1}}\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}) = 0. \tag{12.204}$$

In particular, and since γ^5 is the product of and even number of γ–matrices, we have

$$\text{Tr}\left(\gamma^\mu\gamma^5\right) = i\text{Tr}\,(\gamma^\mu\gamma_0\gamma_1\gamma_2\gamma_3) = 0, \tag{12.205}$$

and

$$\text{Tr}\left(\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_3}\cdots\gamma^{\mu_{2n-1}}\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\gamma^5\right) = 0. \tag{12.206}$$

Similarly, and as shown in Example 12.5 and as illustrated in Problem (12.16) on Page 841, the trace of the product of an odd number of slashed 4-vectors is also zero

$$\text{Tr}\,(\not{a}) = \text{Tr}\left(\not{a}\not{b}\not{c}\right) = \text{Tr}\left(\not{a}\not{b}\not{c}\not{d}\not{e}\right) = \text{Tr}\left(\not{a}\not{b}\not{c}\not{d}\not{e}\not{f}\not{g}\right) = \cdots = 0; \tag{12.207}$$

more generally, the trace of $(2n+1)$ slashed 4-vectors, $(\not{a}_1\not{a}_2\not{a}_3\cdots\not{a}_{2n-1}\not{a}_{2n}\not{a}_{2n+1})$, is zero:

$$\text{Tr}\,(\not{a}_1\not{a}_2\not{a}_3\cdots\not{a}_{2n-1}\not{a}_{2n}\not{a}_{2n+1}) = 0. \tag{12.208}$$

We can also show the following theorems:

$$\text{Tr}\,(\gamma^\mu\gamma^\nu) = 4g^{\mu\nu}, \qquad \text{Tr}\left(\gamma^\mu\gamma^\nu\gamma^5\right) = 0, \qquad \text{Tr}\left(\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^5\right) = 0, \tag{12.209}$$

[27] As a reminder, we have shown in Chapter 2 that trace of the product of a number of operators is invariant under cyclic permutations of these operators: $\text{Tr}\left(\hat{A}\hat{B}\hat{C}\hat{D}\cdots\right) = \text{Tr}\left(\hat{B}\hat{C}\hat{D}\cdots\hat{A}\right) = \text{Tr}\left(\hat{C}\hat{D}\cdots\hat{A}\hat{B}\right) = \text{Tr}\left(\hat{D}\cdots\hat{A}\hat{B}\hat{C}\right) = \cdots$.

$$\mathrm{Tr}\left(\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\alpha\right) = 4\left(g^{\mu\nu}g^{\lambda\alpha} - g^{\mu\lambda}g^{\nu\alpha} + g^{\mu\alpha}g^{\nu\lambda}\right), \tag{12.210}$$

$$\mathrm{Tr}\left(\not{a}\not{b}\right) = 4a\cdot b, \qquad \mathrm{Tr}\left(\gamma^5\not{a}\not{b}\right) = 0, \tag{12.211}$$

$$\mathrm{Tr}\left(\not{a}\not{b}\not{c}\not{d}\right) = 4\left[(a\cdot b)(c\cdot d) - (a\cdot c)(b\cdot d) + (a\cdot d)(b\cdot c)\right]. \tag{12.212}$$

Example 12.4
Using the algebraic properties of the γ–matrices, show the following two relations:
(a) $\gamma^\mu\gamma_\mu = 4I_4$
(b) $\gamma^\nu\gamma^\mu\gamma_\nu = -2\gamma^\mu$.

Solution
(a) First, since $\gamma_\mu = g_{\mu\nu}\gamma^\nu$, we can write

$$\gamma^\mu\gamma_\mu = \gamma^\mu g_{\mu\nu}\gamma^\nu = g_{\mu\nu}\gamma^\mu\gamma^\nu = \frac{1}{2}\left(g_{\mu\nu} + g_{\nu\mu}\right)\gamma^\mu\gamma^\nu, \tag{12.213}$$

where, in the last step, we have used the fact that the Minkowski metric is symmetric, $g_{\mu\nu} = g_{\nu\mu}$. Since μ and ν are dummy labels, we can relabel the last term of (12.213) as $g_{\nu\mu}\gamma^\mu\gamma^\nu \equiv g_{\mu\nu}\gamma^\nu\gamma^\mu$ and rewrite (12.213) in the following form

$$\gamma^\mu\gamma_\mu = \frac{1}{2}g_{\mu\nu}\gamma^\mu\gamma^\nu + \frac{1}{2}g_{\nu\mu}\gamma^\mu\gamma^\nu = \frac{1}{2}g_{\mu\nu}\gamma^\mu\gamma^\nu + \frac{1}{2}g_{\mu\nu}\gamma^\nu\gamma^\mu = \frac{1}{2}g_{\mu\nu}\left(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu\right). \tag{12.214}$$

Now, using the fact that $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4$ as displayed in Eq. (12.181) and since $g_{\mu\nu}g^{\mu\nu} = 4$, we can recast (12.214) in the following form:

$$\gamma^\mu\gamma_\mu = \frac{1}{2}g_{\mu\nu}\left(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu\right) = \frac{1}{2}g_{\mu\nu}\left(2g^{\mu\nu}I_4\right) = g_{\mu\nu}g^{\mu\nu}I_4 = 4I_4. \tag{12.215}$$

(b) Using the relation $\gamma^\mu\gamma_\nu = 2g^\mu_\nu I_4 - \gamma_\nu\gamma^\mu$, which is obtained from Eq. (12.181), and multiplying it by γ^ν on the left, we have

$$\begin{aligned}\gamma^\nu\gamma^\mu\gamma_\nu &= \gamma^\nu\left(2g^\mu_\nu I_4 - \gamma_\nu\gamma^\mu\right) = 2\gamma^\nu g^\mu_\nu I_4 - \gamma^\nu\gamma_\nu\gamma^\mu = 2\gamma^\mu - 4\gamma^\mu \\ &= -2\gamma^\mu,\end{aligned} \tag{12.216}$$

Where we have used the fact that $\gamma^\nu g^\mu_\nu = \gamma^\mu$ and $\gamma^\nu\gamma_\nu = 4$ (see equation (12.184)).

Example 12.5
(a) Show that the trace of the product of an odd number of γ matrices is zero.
(b) Show the relations: $\mathrm{Tr}\,(\not{a}) = 0$ and $\mathrm{Tr}\left(\not{a}\not{b}\not{c}\right) = 0$ where $\not{a}$, $\not{b}$, and $\not{c}$ are slashed 4-vectors.
(c) Use the result of (b) to infer the fact that the trace of the product of an odd number of slashed 4-vectors is zero.

Solution

(a) Using the relation $\left(\gamma^5\right)^2 = 1$ along with the fact that the trace of the product of a number of matrices is invariant under their *cyclic permutations* (see Eq. (12.203)), we can write the product of an *odd number*, $(2n+1)$, of γ matrices as follows:

$$\begin{aligned}\text{Tr}\left(\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\right) &= \text{Tr}\left(\gamma^5\gamma^5\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\right)\\ &= \text{Tr}\left(\gamma^5\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\gamma^5\right).\end{aligned} \tag{12.217}$$

Now, since $\gamma^5\gamma^\mu = -\gamma^\mu\gamma^5$, we can execute $(2n+1)$ permutations of the left-side γ^5 in the trace with all the γ matrices until we reach the right-side γ^5 inside the trace above:

$$\begin{aligned}\text{Tr}\left(\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\right) &= \text{Tr}\left(\gamma^5\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\gamma^5\right)\\ &= -\text{Tr}\left(\gamma^{\mu_1}\gamma^5\gamma^{\mu_2}\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\gamma^5\right)\\ &= \text{Tr}\left(\gamma^{\mu_1}\gamma^{\mu_2}\gamma^5\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\gamma^5\right)\\ &= (-1)^{2n+1}\text{Tr}\left(\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\gamma^5\gamma^5\right)\\ &= -\text{Tr}\left(\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\gamma^5\gamma^5\right)\\ &= -\text{Tr}\left(\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\right)\\ &= 0,\end{aligned} \tag{12.218}$$

where we have used the relation $\gamma^5\gamma^5 = 1$ in the sixth line.

(b) The proof of $\text{Tr}\left(\not{a}\right) = 0$ is quite straightforward:

$$\text{Tr}\left(\not{a}\right) = \text{Tr}\left(a_\mu\gamma^\mu\right) = a_\mu\text{Tr}\left(\gamma^\mu\right) = 0, \tag{12.219}$$

where we have used the facts that a_μ is a 4-vector which can be pulled out of the trace; hence, the trace of γ^μ is zero, since $\text{Tr}\left(\gamma^0\right) = \text{Tr}\left(\gamma^1\right) = \text{Tr}\left(\gamma^2\right) = \text{Tr}\left(\gamma^3\right) = 0$.

Using the same method above, we can easily obtain:

$$\text{Tr}\left(\not{a}\not{b}\not{c}\right) = \text{Tr}\left(a_\mu b_\nu c_\lambda\gamma^\mu\gamma^\nu\gamma^\lambda\right) = a_\mu b_\nu c_\lambda\text{Tr}\left(\gamma^\mu\gamma^\nu\gamma^\lambda\right) = 0, \tag{12.220}$$

where we have used the theorem listed in Eq. (12.204); namely, the trace of the product of an *odd number* of γ^μ matrices is *zero*.

(c) Using (12.218), we can easily generalize the relation (12.220) to the the product of an odd number of slashed 4-vectors:

$$\text{Tr}\left(\not{a}_1\not{a}_2\not{a}_3\cdots\not{a}_{2n-1}\not{a}_{2n}\not{a}_{2n+1}\right) = a_1a_2a_3\cdots a_{2n-1}a_{2n}a_{2n+1}\text{Tr}\left(\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2n}}\gamma^{\mu_{2n+1}}\right) = 0. \tag{12.221}$$

12.4.4 Dirac, Weyl, and Majorana Representations of the γ Matrices

The algebraic properties discussed above are independent of the specific representation of the γ–matrices; sometimes, however, we need to specify explicit representations of the γ–matrices to carry out specific calculations. In this subsection, we are going to look at the three most

common representations: the Dirac, Weyl, and Majorana representations. It is important to note that no matter which representation we chose, the four γ–matrices (i.e., $\gamma^0, \gamma^1, \gamma^2, \gamma^3$) and the fifth matrix γ^5 must obey the Clifford algebra (12.181). In most situations, however, we don't have to use any of the explicit representations; we can work directly with the Clifford algebra.

12.4.4.1 Dirac's Standard Representation of the γ Matrices

In the *standard representation*, also called the *Dirac representation*, the γ^μ matrices are expressed in terms of Dirac's β and α_j matrices (listed in equations (12.161) and (12.163)) as follows:

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \tag{12.222}$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \qquad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{12.223}$$

Taking the product of these matrices, we can verify that the matrix representing γ^5 is given by

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{12.224}$$

Notice that we can rewrite the matrices (12.222) – (12.224) in terms of the 2×2 Pauli matrices σ_j as follows:

$$\gamma^0 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}, \qquad \gamma^j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix} \implies \gamma^5 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}, \qquad j = x, y, z, \tag{12.225}$$

where I_2 is the 2×2 unity matrix and σ_j are listed in (12.162). The matrices (12.222) – (12.224), or their condensed version (12.225), are known as the *Dirac representation* or the *standard representation* of the γ matrices. The set of matrices $\{\gamma^0, \gamma^1, \gamma^2, \gamma^3\}$ is known as the *Dirac basis* for the Clifford algebra. In this context, we can easily verify that the matrices (12.154) – (12.155) satisfy the anticommutation relation (12.181).

12.4.4.2 The Weyl and Majorana Representations of the γ Matrices

In addition to the Dirac basis, we may mention two other widely used bases for the Clifford algebra, the Weyl and the Majorana bases:

- **Weyl or chiral basis:**[28] the *Weyl or chiral representation* uses a basis in which $\gamma^1, \gamma^2, \gamma^3$ are similar to those of (12.225), but γ^0 is different; hence, γ^5 is also different. In this representation, γ^5 is diagonal:

$$\gamma^0 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}, \qquad \gamma^j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix} \implies \gamma^5 = \begin{pmatrix} -I_2 & 0 \\ 0 & I_2 \end{pmatrix}, \qquad j = x, y, z. \tag{12.226}$$

[28] The meaning of "chiral" will be explained in Subsection 12.4.18.

- **Mojarana basis:** The other most common representation is the *Majorana representation* in which all the γ matrices are imaginary so as to give us the same metric $(+,-,-,-)$ as (12.18):

$$\gamma^0 = \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix}, \gamma^1 = \begin{pmatrix} i\sigma_z & 0 \\ 0 & i\sigma_z \end{pmatrix}, \gamma^2 = \begin{pmatrix} 0 & -\sigma_y \\ \sigma_y & 0 \end{pmatrix}, \gamma^3 = \begin{pmatrix} -i\sigma_x & 0 \\ 0 & -i\sigma_x \end{pmatrix} \Longrightarrow \gamma^5 = \begin{pmatrix} \sigma_y & 0 \\ 0 & -\sigma_y \end{pmatrix}. \tag{12.227}$$

  Unlike the Dirac matrices where only γ^2 is imaginary, all the Majorana matrices are imaginary, $(\gamma^\mu)^* = -\gamma^\mu$. Hence, using the Majorana basis, we can easily verify that the spinors that are solutions of the Dirac equation are going to be real: $\psi^* = \psi$; these are called the *Majorana spinors*.

It is important to mention that, when the energy of the particle is too low for the rest-mass energy to be neglected (i.e., for massive particles with low kinetic energies), the Dirac representation is most convenient; electrons, for instance, are prime candidates for Dirac particles. However, in the limit of ultra-high energies (i.e., for massless particles), the Weyl representation is more appropriate; since neutrinos were considered to be massless for many decades, they were initially represented by Weyl spinors (see Subsection 12.4.18).

12.4.5 Dirac Equation in Covariant Form

First, let us express the Dirac Hamiltonian (12.147) in terms of the γ–matrices. Since $\beta = \gamma^0$, $(\gamma^0)^2 = I_4$, and since $\vec{\alpha} \equiv \gamma^0(\gamma^0\vec{\alpha}) = \gamma^0\vec{\gamma}$, we can rewrite the Dirac Hamiltonian as

$$\hat{H} = c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2 \equiv c\gamma^0\vec{\gamma}\cdot\hat{\vec{p}} + \gamma^0 mc^2. \tag{12.228}$$

Now, to determine the covariant form of the Dirac equation, let us begin by multiplying both sides of (12.150) by β/c; this leads to

$$i\hbar\frac{\beta}{c}\frac{\partial}{\partial t}\psi(x^\mu) = \left(-i\hbar\beta\vec{\alpha}\cdot\vec{\nabla} + \beta^2 mc\right)\psi(x^\mu), \tag{12.229}$$

which, when combined with $\beta = \gamma^0$, $\beta\vec{\alpha} = \vec{\gamma}$ and $\beta^2 = 1$, becomes

$$i\hbar\frac{\gamma^0}{c}\frac{\partial}{\partial t}\psi(x^\mu) = \left(-i\hbar\vec{\gamma}\cdot\vec{\nabla} + mc\right)\psi(x^\mu). \tag{12.230}$$

Using Equation (12.197), we can reduce (12.230) to

$$\boxed{\left(i\hbar\gamma^\mu\partial_\mu - mc\right)\psi(x^\mu) = 0 \quad \longrightarrow \quad \left(i\hbar\partial\!\!\!/ - mc\right)\psi(x^\mu) = 0.} \tag{12.231}$$

This is the *Dirac equation in covariant form* for a free electron; notice that the time and space derivatives are treated alike, both are first order. Dirac[29] introduced this equation in 1928. Additionally, using (12.198), we can express (12.231) in terms of the 4-momentum operator, $\hat{p}^\mu = i\hbar\partial^\mu$, as follows

$$\boxed{\left(\gamma^\mu\hat{p}_\mu - mc\right)\psi(x^\mu) = 0 \quad \longrightarrow \quad \left(\hat{\not{p}} - mc\right)\psi(x^\mu) = 0,} \tag{12.232}$$

where $\hat{\not{p}} = \gamma^\mu\hat{p}_\mu$.

[29] P.A.M. Dirac, *Proc. Roy. Soc.* (London) **A117**, 610 (1928).

12.4.6 Lorentz Covariance of the Dirac Equation

Our aim here is to show that the Dirac equation is *form-invariant* under Lorentz transformations; that is, we want to show that the form of the Dirac equation is the same in all inertial frames – frames that are connected by Lorentz transformations.

Consider two inertial frames F and F' and two observers, O in F and O' in F'. Let us assume that O observes an event taking place at x^μ in F; O' will instead observe that same event taking place at x'^μ in F' ; the transformation from x^μ to x'^μ is given by:

$$x^\mu \longrightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu, \tag{12.233}$$

where Λ is a Lorentz transformation; its coefficients $\Lambda^\mu{}_\nu$ depend on the relative velocities and spatial orientations of F and F'. For instance, one might consider the Lorentz boost along the x–axis listed above; see equation (12.9). As discussed above, the "length" $x_\mu x^\mu$ is conserved under Lorentz transformations, $x_\mu x^\mu = x'_\mu x'^\mu$. Using

$$x_\mu x^\mu = g_{\mu\nu}\, x^\mu x^\mu, \tag{12.234}$$

and

$$x'_\mu x'^\mu = g_{\mu\nu}\, x'^\mu\, x'^\mu = g_{\mu\nu}\, \Lambda^\mu{}_\alpha\, x^\alpha \Lambda^\mu{}_\beta\, x^\beta \equiv g_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\mu{}_\beta\, x^\alpha x^\beta, \tag{12.235}$$

we obtain

$$x_\mu x^\mu = x'_\mu x'^\mu \quad \Longrightarrow \quad g_{\alpha\beta} = g_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\mu{}_\beta \quad \Longrightarrow \quad \Lambda^\mu{}_\alpha \Lambda^\alpha{}_\nu = \delta^\mu_\nu = g^\mu_\nu. \tag{12.236}$$

Using (12.233), we can write the Lorentz transformation of 4-derivatives ∂_ν as follows:

$$\frac{\partial}{\partial x^\nu} = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial}{\partial x'^\mu} = \Lambda^\mu{}_\nu \frac{\partial}{\partial x'^\mu} \quad \longrightarrow \quad \partial_\nu = \Lambda^\mu{}_\nu \partial'_\mu \quad \Longleftrightarrow \quad \partial'_\mu = [\Lambda^{-1}]^\nu{}_\mu\, \partial_\nu. \tag{12.237}$$

12.4.6.1 Lorentz Transformations of the Dirac Equation

Let us assume that Observer O describes a particle in the frame F by the spinor $\psi(x^\mu)$ which obeys the Dirac equation:

$$\left(i\hbar\gamma^\mu\partial_\mu - mc\right)\psi(x^\mu) = 0. \tag{12.238}$$

The important issue now is to answer this question: Which Dirac equation will Observer O' utilize to describe the same particle in F'? If the Dirac equation in F' will have the same form as (12.238), then the Dirac equation is *Lorentz covariant*. This issue reduces to finding the symmetry transformation $S(\Lambda)$ that would transform the spinor $\psi(x)$ into $\psi'(x')$:

$$\boxed{\psi(x^\mu) \longrightarrow \psi'(x'^\mu) = S(\Lambda)\psi(x^\mu) \qquad \Longrightarrow \qquad \psi(x^\mu) = S^{-1}\psi'(x'^\mu),} \tag{12.239}$$

where $S(\Lambda)$ can be represented by a 4×4 matrix whose coefficients depend on the Lorentz transformation $\Lambda^\mu{}_\nu$; for instance, we may consider a symmetry transformation that is related to a boost: $\psi' = S_{Boost}(\beta)\psi$ where $\beta = v/c$. Inserting $\psi(x^\mu) = S^{-1}\psi'(x'^\mu)$ into the Dirac equation (12.238) and multiplying on the left by S, we obtain

$$\left(i\hbar S\gamma^\mu S^{-1}\partial_\mu - mcSS^{-1}\right)\psi'(x'^\mu) = 0 \quad \Longrightarrow \quad \left(i\hbar S\gamma^\mu S^{-1}\partial_\mu - mc\right)\psi'(x'^\mu) = 0, \tag{12.240}$$

where we have used the fact that $SS^{-1} = I_4$. From (12.237), we have $\partial_\mu = \Lambda^\nu{}_\mu \partial'_\nu$, which, when inserted into (12.240), leads to the Dirac equation in F':

$$\left(i\hbar S\gamma^\mu S^{-1}\Lambda^\nu{}_\mu \partial'_\nu - mc\right)\psi'(x'^\nu) = 0. \tag{12.241}$$

A comparison of (12.238) and (12.241) leads inevitably to the following conclusion: in order for the Dirac equation (12.241) to have the same form as the original equation (12.238) or, in order for the Dirac equation to be form invariant under Lorentz transformations, $S(\Lambda)$ has to satisfy the following condition:

$$\boxed{S\gamma^\mu S^{-1}\Lambda^\nu{}_\mu = \gamma^\nu \quad \Longleftrightarrow \quad S\gamma^\mu S^{-1} = [\Lambda^{-1}]^\mu{}_\nu \gamma^\nu \quad \text{or} \quad \Lambda^\nu{}_\mu \gamma^\mu = S^{-1}\gamma^\nu S.} \tag{12.242}$$

If this condition is satisfied, the Dirac equation (12.241) reduces to

$$(i\hbar\gamma^\nu \partial'_\nu - mc)\,\psi'(x'^\nu) = 0. \tag{12.243}$$

This is the Dirac equation in F' which is obtained from a transformation of (12.238). Notice that, after transforming (12.238) into (12.243), ∂_μ and $\psi(x^\mu)$ have changed but not γ^μ; this was expected since the $\gamma-$matrices are constant and, hence, they are the same in all inertial frames; that is, they are invariant under Lorentz transformations: $\gamma'^\nu = \gamma^\nu$.

12.4.6.2 Determination of $S(\Lambda)$ Using Infinitesimal Lorentz Transformations

Now, the real issue boils down to solving equation (12.242) and deriving the symmetry transformation $S(\Lambda)$. To construct $S(\Lambda)$, we consider an infinitesimal proper Lorentz transformation

$$\Lambda_{\mu\nu} = g_{\mu\nu} + \varepsilon_{\mu\nu}, \tag{12.244}$$

where where $[\varepsilon]$ is an infinitesimal antisymmetric matrix: $\varepsilon_{\mu\nu} = -\varepsilon_{\nu\mu}$. This can be seen as follows: Inserting (12.244) into the relation $g_{\alpha\beta} = \Lambda^\mu{}_\alpha \Lambda_{\mu\beta}$ (which can be obtained from (12.236)) and keeping only the first order terms in $\varepsilon_{\beta\alpha}$, we obtain

$$\begin{aligned} g_{\alpha\beta} &= \Lambda^\mu{}_\alpha \Lambda_{\mu\beta} = \left(g^\mu{}_\alpha + \varepsilon^\mu{}_\alpha\right)\left(g_{\mu\beta} + \varepsilon_{\mu\beta}\right) \\ &= g^\mu{}_\alpha g_{\mu\beta} + g^\mu{}_\alpha \varepsilon_{\mu\beta} + \varepsilon^\mu{}_\alpha g_{\mu\beta} \\ &= g_{\alpha\beta} + \varepsilon_{\alpha\beta} + \varepsilon_{\beta\alpha}, \end{aligned} \tag{12.245}$$

and hence: $\varepsilon_{\alpha\beta} = -\varepsilon_{\beta\alpha}$. Thus, $[\varepsilon]$ can have only 6 independent nonvanishing elements.

Let us now expand $S(\Lambda)$ to first order in $\varepsilon_{\mu\nu}$:

$$S = I_4 - \frac{i}{4}\varepsilon_{\mu\nu}\sigma^{\mu\nu}, \tag{12.246}$$

and hence

$$S^{-1} = I_4 + \frac{i}{4}\varepsilon_{\mu\nu}\sigma^{\mu\nu}, \tag{12.247}$$

where $\sigma^{\mu\nu}$ is a 4×4 matrix. Note that, since $\varepsilon_{\mu\nu}$ is antisymmetric, $\varepsilon_{\mu\nu} = -\varepsilon_{\nu\mu}$, we can antisymmetrize $\varepsilon_{\mu\nu}\sigma^{\mu\nu}$ as follows (regardless of the symmetry nature of $\sigma^{\mu\nu}$):

$$\begin{aligned}\varepsilon_{\mu\nu}\sigma^{\mu\nu} &= \frac{1}{2}\left(\varepsilon_{\mu\nu}\sigma^{\mu\nu} + \varepsilon_{\mu\nu}\sigma^{\mu\nu}\right) = \frac{1}{2}\left(\varepsilon_{\mu\nu}\sigma^{\mu\nu} - \varepsilon_{\nu\mu}\sigma^{\mu\nu}\right) \\ &= \frac{1}{2}\left(\varepsilon_{\mu\nu}\sigma^{\mu\nu} - \varepsilon_{\mu\nu}\sigma^{\nu\mu}\right) \\ &= \frac{1}{2}\,\varepsilon_{\mu\nu}\,(\sigma^{\mu\nu} - \sigma^{\nu\mu}),\end{aligned} \tag{12.248}$$

where, in the second line, we have switched the μ and ν labels: $\mu \longleftrightarrow \nu$. As we will see below (see Eq. (12.254)), $\sigma^{\mu\nu}$ is antisymmetric: $\sigma^{\mu\nu} = -\sigma^{\nu\mu}$.

Inserting (12.246) and (12.247) into the condition 12.242), we obtain

$$\Lambda^{\nu}{}_{\mu}\gamma^{\mu} = S^{-1}\gamma^{\nu}S \quad\Longrightarrow\quad \left(g^{\nu}{}_{\mu} + \varepsilon^{\nu}{}_{\mu}\right)\gamma^{\mu} = \left(I_4 + \frac{i}{4}\varepsilon_{\alpha\beta}\sigma^{\alpha\beta}\right)\gamma^{\nu}\left(I_4 - \frac{i}{4}\varepsilon_{\alpha\beta}\sigma^{\alpha\beta}\right). \tag{12.249}$$

An expansion of the right-hand side of the last relation to first order in $\sigma^{\mu\nu}$ leads to

$$\gamma^{\nu} + \varepsilon^{\nu}{}_{\mu}\gamma^{\mu} = \gamma^{\nu} - \frac{i}{4}\varepsilon_{\alpha\beta}\,\gamma^{\nu}\,\sigma^{\alpha\beta} + \frac{i}{4}\varepsilon_{\alpha\beta}\sigma^{\alpha\beta}\,\gamma^{\nu}, \tag{12.250}$$

which reduces to

$$\varepsilon^{\nu}{}_{\mu}\gamma^{\mu} = -\frac{i}{4}\varepsilon_{\alpha\beta}\left[\gamma^{\nu},\,\sigma^{\alpha\beta}\right]. \tag{12.251}$$

For notational consistency between both sides of the above equation, let us express the left-hand side in terms $\varepsilon_{\alpha\beta}$. Since $\varepsilon^{\nu}{}_{\mu}\gamma^{\mu} \equiv \varepsilon^{\nu}{}_{\beta}\gamma^{\beta}$ and following the same method used in equation (12.248), we can write

$$\varepsilon^{\nu}{}_{\beta}\gamma^{\beta} = \varepsilon_{\alpha\beta}g^{\nu\alpha}\gamma^{\beta} = \frac{1}{2}\varepsilon_{\alpha\beta}\left(g^{\nu\alpha}\gamma^{\beta} - g^{\nu\beta}\gamma^{\alpha}\right). \tag{12.252}$$

Inserting (12.252) into the left-hand side of (12.251), we obtain

$$\frac{1}{2}\varepsilon_{\alpha\beta}\left(g^{\nu\alpha}\gamma^{\beta} - g^{\nu\beta}\gamma^{\alpha}\right) = -\frac{i}{4}\varepsilon_{\alpha\beta}\left[\gamma^{\nu},\,\sigma^{\alpha\beta}\right] \quad\Longrightarrow\quad 2i\left(g^{\nu\alpha}\gamma^{\beta} - g^{\nu\beta}\gamma^{\alpha}\right) = \left[\gamma^{\nu},\,\sigma^{\alpha\beta}\right]. \tag{12.253}$$

Hence, the problem of proving the Lorentz covariance of the Dirac equation is now reduced to finding the six independent elements of $\sigma^{\alpha\beta}$ since, as mentioned above, $[\sigma]$ is a 4×4 antisymnetric matrix. We may get a hint from of equation (12.181) which involves the product of $\gamma-$matrices. Moreover, the construction of $\sigma^{\alpha\beta}$ in terms of the $\gamma-$matrices was introduced in equation (12.186); that is

$$\sigma^{\mu\nu} = \frac{i}{2}\left[\gamma^{\mu},\,\gamma^{\nu}\right]. \tag{12.254}$$

Finally, inserting (12.254) into (12.246), we obtain the expression for the symmetry transformation S that corresponds to an infinitesimal Lorentz transforation:

$$\boxed{S = I_4 - \tfrac{i}{4}\varepsilon_{\mu\nu}\sigma^{\mu\nu} = I_4 + \tfrac{1}{8}\varepsilon_{\mu\nu}\left[\gamma^{\mu},\,\gamma^{\nu}\right].} \tag{12.255}$$

So, the Dirac equation is Lorentz covariant under this infinitesimal transformation.

12.4.6.3 Determination of $S(\Lambda)$ Using Finite Lorentz Transformations

Let us now focus on how to constructing $S(\Lambda)$ for *finite* Lorentz transformation Λ^{μ}_{ν}. Instead of the infinitesimal transformation (12.255), one can construct S for a finite transformation by a succession of a large number of infinitesimal transformations. In practice, consider we can divide the finite transformation "angle" or parameter $\alpha_{\mu\nu}$ into N small parameters $\varepsilon_{\mu\nu}$; that is: $\alpha_{\mu\nu} = N\varepsilon_{\mu\nu}$. So, the transformation over the finite parameter $\varepsilon_{\mu\nu}$ can thus be viewed as a series of N consecutive infinitesimal transformations each given by equation (12.246): $S_\varepsilon(\Lambda) = I_4 - \frac{i}{4}\varepsilon_{\mu\nu}\sigma^{\mu\nu}$. Hence, we can write express the finite transformation $S(\Lambda(\alpha_{\mu\nu}))$ in terms of N infinitesimal transformations $S_\varepsilon(\Lambda)$ as follows:

$$S(\Lambda)) = S_\varepsilon(\Lambda)S_\varepsilon(\Lambda)\cdots S_\varepsilon(\Lambda) = [S_\varepsilon(\Lambda)]^N = \left(1 - \frac{i}{4}\frac{\alpha_{\mu\nu}}{N}\sigma^{\mu\nu}\right)^N. \tag{12.256}$$

Since $\varepsilon_{\mu\nu} = \alpha_{\mu\nu}/N$, and if $\varepsilon_{\mu\nu}$ is infinitesimally small, we have

$$S = \lim_{N\to\infty}\left(1 - \frac{i}{4}\frac{\alpha_{\mu\nu}}{N}\sigma^{\mu\nu}\right)^N, \tag{12.257}$$

which, when combined with the well-known expression $\lim_{N\to\infty}(1 + x/N) = e^x$, leads to

$$\boxed{S = \exp\left(-\tfrac{i}{4}\alpha_{\mu\nu}\sigma^{\mu\nu}\right).} \tag{12.258}$$

So, under this finite transformation, the Dirac spinor transforms as follows:

$$\psi'(x') = \exp\left(-\frac{i}{4}\alpha_{\mu\nu}\sigma^{\mu\nu}\right)\psi(x). \tag{12.259}$$

The Dirac equation corresponding to this *finite* transformation is Lorentz covariant.

We should note that $\sigma^{\mu\nu}$ is the generator of the S transformation; it can be the generator of rotations and boosts in the Dirac space. For instance, we can show that

$$\sigma^{0j} = \frac{i}{2}\left[\gamma^0, \gamma^j\right] = i\gamma^0\gamma^j = i\begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix}, \qquad j = 1, 2, 3, \tag{12.260}$$

and

$$\sigma^{jk} = \frac{i}{2}\left[\gamma^j, \gamma^k\right] = i\gamma^j\gamma^k = -i\begin{pmatrix} [\sigma_j, \sigma_k] & 0 \\ 0 & [\sigma_j, \sigma_k] \end{pmatrix} = 2\epsilon_{jkl}\begin{pmatrix} \sigma_l & 0 \\ 0 & \sigma_l \end{pmatrix} =, \qquad j, k, l = 1, 2, 3, \tag{12.261}$$

where σ_j is the 2×2 Pauli matrix along the $j-$axis and where we have used the fact that $[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l$. Clearly, the matrix σ^{0j} is the generator of a boost and σ^{jk} the generator of a rotation. As an illustration of such finite transformations, we consider the following two cases: a finite boost transformation along the $x-$axis and a finite rotation around the $z-$axis:

- Boost along the $x-$axis by a finite parameter χ: Since $\sigma^{01} = i\begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}$ and $\alpha_{01} = -\alpha_{10} = \chi$, we have shown in Problem 12.18 that the transformation (12.258) reduces to

$$\boxed{S = \exp\left(-\tfrac{i}{2}\chi\sigma^{01}\right) = \exp\left[\tfrac{1}{2}\chi\begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}\right] = \begin{pmatrix} I_2 & 0 \\ 0 & I_2 \end{pmatrix}\cosh\left(\tfrac{\chi}{2}\right) + \begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}\sinh\left(\tfrac{\chi}{2}\right),} \tag{12.262}$$

where I_2 and I_4 are the 2×2 and 4×4 unity matrices, respectively, and where σ_1 is the 2×2 Pauli matrix along the x–axis: $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

- Rotation about the z–axis by a finite angle ϕ: Since $\sigma^{12} = \begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}$ and $\alpha_{21} = -\alpha_{12} = \phi$, we have shown in Problem 12.17 that the transformation (12.258) reduces to

$$\boxed{S = \exp\left(-\tfrac{i}{2}\phi\sigma^{12}\right) = \exp\left[-\tfrac{i}{2}\phi\begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}\right] = \begin{pmatrix} I_2 & 0 \\ 0 & I_2 \end{pmatrix}\cos\left(\tfrac{\phi}{2}\right) - i\begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}\sin\left(\tfrac{\phi}{2}\right),} \tag{12.263}$$

where σ_3 is the 2×2 Pauli matrix along the z–axis: $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Clearly, the Dirac spinor (12.259) would transform under $S = \exp\left[-\tfrac{i}{2}\phi\begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}\right]$ like a spin-1/2 particle.

12.4.7 Continuity Equation and Current Conservation

We are going to follow two approaches to derive the current conservation by starting first from the non-covariant form of the Dirac equation (12.150) and, second, by starting from the covariant form of the Dirac equation (12.231).

First approach: Starting from the non-covariant form of the Dirac equation

Since the matrices of α_j and β are Hermitian and since $\vec{\nabla}$ is anti-Hermitian (as shown in Chapter 3), we can write the Hermitian conjugate of (12.150) as follows

$$-i\hbar\frac{\partial}{\partial t}\psi^\dagger(x^\mu) = \left[i\hbar c\left(\vec{\nabla}\psi^\dagger(x^\mu)\right)\cdot\vec{\alpha} + mc^2\psi^\dagger(x^\mu)\beta\right]. \tag{12.264}$$

Multiplying (12.150) on the left by $\psi^\dagger(x^\mu)$ and (12.264) on the right by $\psi(x^\mu)$, and subtracting the two resulting equations (i.e., $\psi^\dagger(x^\mu)\times$(12.150) – (12.264)$\times\psi(x^\mu)$), we obtain[30]

$$\begin{aligned} i\hbar\psi^\dagger\frac{\partial\psi}{\partial t} + i\hbar\frac{\partial\psi^\dagger}{\partial t}\psi \;=\; & \psi^\dagger\left[-i\hbar c\vec{\alpha}\cdot\left(\vec{\nabla}\psi\right) + mc^2\beta\psi\right] \\ & -\left[i\hbar c\left(\vec{\nabla}\psi^\dagger\right)\cdot\vec{\alpha} + mc^2\psi^\dagger\beta\right]\psi, \end{aligned} \tag{12.265}$$

which reduces to

$$i\hbar\frac{\partial}{\partial t}\left[\psi^\dagger\psi\right] = -i\hbar c\psi^\dagger\left[\vec{\alpha}\cdot\left(\vec{\nabla}\psi\right)\right] - i\hbar c\left[\left(\vec{\nabla}\psi^\dagger\right)\cdot\vec{\alpha}\right]\psi. \tag{12.266}$$

Using the identity

$$\psi^\dagger\left[\vec{\alpha}\cdot\left(\vec{\nabla}\psi\right)\right] + \left[\left(\vec{\nabla}\psi^\dagger\right)\cdot\vec{\alpha}\right]\psi = \vec{\nabla}\cdot\left(\psi^\dagger\vec{\alpha}\psi\right), \tag{12.267}$$

we can reduce (12.266) to

$$\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) + c\vec{\nabla}\cdot\left(\psi^\dagger\vec{\alpha}\psi\right) = 0, \tag{12.268}$$

[30]For notational simplicity, we are going to drop (x^μ) in $\psi^\dagger(x^\mu)$ and $\psi(x^\mu)$ and denote them by $\psi^\dagger$ and ψ, respectively.

or to

$$\boxed{\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \qquad \text{where} \qquad \rho = \psi^\dagger \psi, \qquad \vec{j} = c\psi^\dagger \vec{\alpha} \psi.} \tag{12.269}$$

We will deal with the interpretation of ρ and $\vec{j}$ later on (see the comments offered right after Equation (12.284)).

Second approach: Starting from the covariant form of the Dirac equation

First, we need to take the hermitian conjugate of (12.231)

$$\begin{aligned} \left[\left(i\hbar\gamma^\mu \partial_\mu - mc \right) \psi \right]^\dagger &= \psi^\dagger \left(i\hbar\gamma^\mu \partial_\mu - mc \right)^\dagger \\ &= \psi^\dagger \left(-i\hbar \left(\gamma^\mu\right)^\dagger \overleftarrow{\partial_\mu} - mc \right) = 0, \end{aligned} \tag{12.270}$$

where we have used the fact that $\left(\partial_\mu \psi\right)^\dagger = \psi^\dagger \overleftarrow{\partial_\mu}$ and where $\overleftarrow{\partial_\mu}$ denotes the 4-differential operator which acts to the left side. Since $\left(\gamma^\mu\right)^\dagger = \gamma^0 \gamma^\mu \gamma^0$ (see (12.176)), we can reduce (12.270) to

$$\left[\left(i\hbar\gamma^\mu \partial_\mu - mc \right) \psi \right]^\dagger = \psi^\dagger \left(-i\hbar\gamma^0 \gamma^\mu \gamma^0 \overleftarrow{\partial_\mu} - mc \right) = 0. \tag{12.271}$$

Multiplying (12.271) from the right by γ^0 and using the fact that $(\gamma^0)^2 = \gamma^0$, we obtain

$$\begin{aligned} \left[\left(i\hbar\gamma^\mu \partial_\mu - mc \right) \psi \right]^\dagger &= \psi^\dagger \left(-i\hbar\gamma^0 \gamma^\mu \left(\gamma^0\right)^2 \overleftarrow{\partial_\mu} - mc\gamma^0 \right) \\ &= \psi^\dagger \gamma^0 \left(-i\hbar\gamma^\mu \overleftarrow{\partial_\mu} - mc \right) = 0, \end{aligned} \tag{12.272}$$

which, using the following notation for the adjoint spinor

$$\overline{\psi} \equiv \psi^\dagger \gamma^0, \tag{12.273}$$

reduces to

$$\overline{\psi} \left(i\hbar\gamma^\mu \overleftarrow{\partial_\mu} + mc \right) = 0, \tag{12.274}$$

which, in turn, can be rewritten as

$$\boxed{i\hbar \left(\partial_\mu \overline{\psi} \right) \gamma^\mu + mc\overline{\psi} = 0.} \tag{12.275}$$

This relation is the hermitian conjugate of the covariant form of the Dirac equation (12.231).

Now, multiplying (12.275) on the right by ψ and (12.231) on the left by $\overline{\psi}$ and adding the two relations, we obtain

$$\left[i\hbar \left(\partial_\mu \overline{\psi} \right) \gamma^\mu + mc\overline{\psi} \right] \psi + \overline{\psi} \left[i\hbar\gamma^\mu \partial_\mu \psi - mc\psi \right] = 0, \tag{12.276}$$

or

$$\left(\partial_\mu \overline{\psi} \right) \gamma^\mu \psi + \overline{\psi} \gamma^\mu \left(\partial_\mu \psi \right) = 0. \tag{12.277}$$

We can rewrite this equation in the following condensed form

$$\partial_\mu \left(\overline{\psi}\gamma^\mu \psi\right) = 0. \tag{12.278}$$

Introducing the 4-current density[31] $j^\mu = c\overline{\psi}\gamma^\mu\psi = c\psi^\dagger\gamma^0\gamma^\mu\psi$, we can interpret (12.278) as the probability continuity equation,

$$\boxed{\partial_\mu j^\mu = 0 \quad \longrightarrow \quad \partial_\mu j^\mu = \frac{\partial \rho}{\partial t} + \vec{\nabla}\cdot\vec{j} = 0,} \tag{12.279}$$

where ρ and $\vec{j}$ are given by

$$\boxed{\rho = \psi^\dagger\psi, \qquad \vec{j} = c\psi^\dagger\vec{\alpha}\psi.} \tag{12.280}$$

We can rewrite the expression of j_μ as

$$\boxed{j^\mu = c\psi^\dagger\gamma^0\gamma^\mu\psi = \left(j^0,\ \vec{j}\right) = \begin{cases} j^0 = c\rho = c\psi^\dagger\gamma^0\gamma^0\psi = c\psi^\dagger\psi \\ \vec{j} = c\psi^\dagger\gamma^0\vec{\gamma}\psi = c\psi^\dagger\gamma^0\gamma^0\vec{\alpha}\psi = c\psi^\dagger\vec{\alpha}\psi \end{cases}} \tag{12.281}$$

where the matrices of the spinor ψ and its hermitian conjugate $\psi^\dagger$ are given by

$$\psi^\dagger = \begin{pmatrix} \psi_1^* & \psi_2^* & \psi_3^* & \psi_4^* \end{pmatrix}, \qquad \psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}. \tag{12.282}$$

Note that the adjoint spinor $\overline{\psi}$ is a matrix with one row and four columns; it can be written as follows:

$$\overline{\psi} = \psi^\dagger\gamma^0 = \begin{pmatrix} \psi_1^* & \psi_2^* & \psi_3^* & \psi_4^* \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \psi_1^* & \psi_2^* & -\psi_3^* & -\psi_4^* \end{pmatrix}. \tag{12.283}$$

We can show that $\overline{\psi}\psi$ transforms like a scalar (invariant) under Lorentz transformations

$$\overline{\psi}\psi = \begin{pmatrix} \psi_1^* & \psi_2^* & -\psi_3^* & -\psi_4^* \end{pmatrix}\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \left|\psi_1^*\right|^2 + \left|\psi_2^*\right|^2 - \left|\psi_3^*\right|^2 - \left|\psi_4^*\right|^2. \tag{12.284}$$

The quantities ρ and $\vec{j}$ in (12.279)–(12.281) clearly represent, respectively, the *probability density* and the *probability current density*, or simply the *current density*. Unlike the Klein–Gordon equation, *the Dirac equation naturally leads to a probability density which is always positive definite*

[31] We have multiplied $\overline{\psi}\gamma^\mu\psi$ by the speed of light c to make the expression of j^μ dimensionally consistent with that of a current density.

(as expected for a probability), since ρ consists of the sum of the squared magnitudes of the four components of the spinor $\psi(x^\mu)$:

$$\rho = \psi^\dagger \psi = \sum_{\alpha=1}^{4} |\psi_\alpha|^2, \tag{12.285}$$

where each of $|\psi_\alpha|^2$ (where $\alpha = 0, 1, 2, 3$) is nonnegative and, hence, we have $\rho \geq 0$. One of the most essential reasons that led to the acceptance of Dirac's equation as the correct prescription for relativistic quantum mechanics is its ability to naturally yield a probability density that is positive definite.

As mentioned in this chapter's introduction, the Dirac equation has resolved the two most serious problems that have plagued the Klein–Gordon equation – the negative probability densities and the proper interpretation of the negative energy solutions – as well as the correct description of the electron's spin from first principles. Now that the Klein–Gordon's sticking issue of the negative probability densities has been resolved, let us turn to the problem of the negative energies. In the following section, we are going to discuss this very issue within the context of a simple system – the free particle.

12.4.8 Solutions of the Dirac Equation for a Free Particle at Rest

In what follows, we are going to derive the *free-particle* solutions and then examine their interpretations. We can obtain sufficient insight about these solutions by considering the simple case of a free particle.

Let us begin by studying the solutions of the Dirac equation for a free particle at *rest*, $\vec{p} = 0$. In this case, the spatial derivatives vanish:

$$\gamma^\mu p_\mu = \frac{i\hbar}{c}\gamma^0 \frac{\partial}{\partial t} - \vec{\gamma}\cdot\vec{p} \equiv \frac{i\hbar}{c}\gamma^0 \frac{\partial}{\partial t} = \gamma^0 p_0. \tag{12.286}$$

That is, in the rest frame of the free particle (of mass m), the Dirac equation (12.232) reduces to

$$\left(\gamma^\mu p_\mu - mc\right)\psi(x^\mu) = 0 \quad \longrightarrow \quad \gamma^0 p_0 \psi(x^\mu) = mc\psi(x^\mu), \tag{12.287}$$

or to[32]

$$\frac{i\hbar}{c}\gamma^0 \frac{\partial|\psi(t)\rangle}{\partial t} = mc|\psi(t)\rangle \quad \longrightarrow \quad \gamma^0 \frac{\partial|\psi(t)\rangle}{\partial t} = -\frac{imc^2}{\hbar}|\psi(t)\rangle. \tag{12.288}$$

Multiplying both sides of (12.288) by γ^0 and using the property $\left(\gamma^0\right)^2 = \gamma^0$, we obtain

$$\frac{\partial|\psi(t)\rangle}{\partial t} = -\frac{imc^2}{\hbar}\gamma^0|\psi(t)\rangle. \tag{12.289}$$

After combining the matrix form (12.282) of the spinor ψ and the matrix representation (12.222) for γ^0, we can easily ascertain that (12.289) consists of four (first-order) differential equations:

$$\frac{\partial}{\partial t}\begin{pmatrix}|\psi_1\rangle\\|\psi_2\rangle\\|\psi_3\rangle\\|\psi_4\rangle\end{pmatrix} = -\frac{imc^2}{\hbar}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}|\psi_1\rangle\\|\psi_2\rangle\\|\psi_3\rangle\\|\psi_4\rangle\end{pmatrix}, \tag{12.290}$$

[32]Here, we are going to use the bracket notation, for it will make the presentation's layout more transparent.

or

$$\frac{\partial}{\partial t}\begin{pmatrix}|\psi_1\rangle\\|\psi_2\rangle\\|\psi_3\rangle\\|\psi_4\rangle\end{pmatrix} = -\frac{imc^2}{\hbar} I_4 \begin{pmatrix}|\psi_1\rangle\\|\psi_2\rangle\\-|\psi_3\rangle\\-|\psi_4\rangle\end{pmatrix}, \tag{12.291}$$

where each of the four states $|\psi_1\rangle$, $|\psi_2\rangle$, $|\psi_3\rangle$, $|\psi_4\rangle$ has four components. We now need to solve these four differential equations for $|\psi_1\rangle$, $|\psi_2\rangle$, $|\psi_3\rangle$, and $|\psi_4\rangle$; actually, we could have obtained these equations much quicker by simply inserting $p_x = p_y = p_z = 0$ into (12.167). Now, we can reduce (12.291) to two sets of two differential equations each:

$$\frac{\partial|\psi_j(t)\rangle}{\partial t} = -\frac{imc^2}{\hbar}|\psi_j(t)\rangle, \qquad j = 1,\ 2, \tag{12.292}$$

and

$$\frac{\partial|\psi_j(t)\rangle}{\partial t} = \frac{imc^2}{\hbar}|\psi_j(t)\rangle, \qquad j = 3,\ 4. \tag{12.293}$$

Notice the sign difference between the previous two equations. The solutions of these equations clearly consist of plane waves

$$|\psi_j(t)\rangle = |u_j(E,\ 0)\rangle e^{-imc^2t/\hbar}, \qquad j = 1,\ 2, \tag{12.294}$$

$$|\psi_j(t)\rangle = |u_j(E,\ 0)\rangle e^{imc^2t/\hbar}, \qquad j = 3,\ 4, \tag{12.295}$$

where each $|u_j(E,\ 0)\rangle$ (with $j = 1,\ 2,\ 3,\ 4$) is a 4-component spinor, since each $|\psi_j\rangle$ has four components. Each one of these spinors depends only on the particle's energy E and its momentum p which is zero in this case; for notational brevity, we will be denoting $|u_j(E,\ 0)\rangle$ by $|u_j\rangle$ in the rest of this subsection. The terms $e^{\pm imc^2t/\hbar}$ account for the oscillations of the wave functions which are caused by the particle's rest energy, mc^2. For the four states $|\psi_1\rangle$, $|\psi_2\rangle$, $|\psi_3\rangle$, $|\psi_4\rangle$ to be orthonormalized (i.e., normalized and orthogonal), we may select $|u_1\rangle$, $|u_2\rangle$, $|u_3\rangle$, $|u_4\rangle$ as follows

$$\boxed{|u_1\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \qquad |u_2\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \qquad |u_3\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \qquad |u_4\rangle = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}.} \tag{12.296}$$

By plugging (12.296) into (12.294) and (12.295), we clearly see that the solutions of (12.291) consist of:

- two degenerate solutions with *positive energy* $E_1 = E_2 \equiv E^{(+)} = mc^2$

$$\boxed{E^{(+)} = mc^2, \qquad |\psi_1(t)\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} e^{-imc^2t/\hbar}, \qquad |\psi_2(t)\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix} e^{-imc^2t/\hbar},} \tag{12.297}$$

- and two other degenerate solutions corresponding to a *negative energy* $E_3 = E_4 \equiv E^{(-)} = -mc^2$

$$\boxed{E^{(-)} = -mc^2, \qquad |\psi_3(t)\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{imc^2 t/\hbar}, \qquad |\psi_4(t)\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{imc^2 t/\hbar}.} \tag{12.298}$$

Again, notice that the signs of the time exponents in $|\psi_3(t)\rangle$ and $|\psi_4(t)\rangle$ are reversed with respect to those of $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$. Like the Klein–Gordon equation, the Dirac equation also encounters negative energy solutions; however, unlike the Klein–Gordon equation, the Dirac equation has a meaningful interpretation for them as will be discussed in Subsections 12.4.10–12.4.12.

The underlying reasons behind the double degeneracy and the negative energy solutions can be viewed, respectively, as follows.

First, the double degeneracy of each energy indicates that these two linearly independent solutions correspond to two different states for the same particle. So we need a new degree of freedom, or an operator, that will provide us with a mechanism to distinguish between these two states. The spin orientation can provide us with such a viable mechanism; the four solutions (12.297)–(12.298) clearly describe two signs of the energy and two possible spin orientations within the particle's rest frame. As our free particle is at rest and if its total spin were 1/2, it can be found in either an up or down spin state, since the projection of a spin 1/2 along an arbitrary axis, say $z-$axis, will be either $+\hbar/2$ or $-\hbar/2$. So, the only meaningful interpretation is that the two degenerate states describe two spin orientations (up and down) of the same particle. Hence, *the Dirac equation describes a spin 1/2 particle where* $|\psi_1(t)\rangle$ *and* $|\psi_2(t)\rangle$ *correspond to spin-up and spin-down configurations; the same applies to* $|\psi_3(t)\rangle$ *and* $|\psi_4(t)\rangle$. This is a quite daring conclusion especially that, up to this point in this chapter, we have not mentioned anything yet about the spin of the particle being described by Dirac's equation; however, there could be little doubt that each doubly degenerate set of linearly independent solutions correspond to two different spin orientations of the particle. In this context, and as shown in Problem (12.20) on Page 850, the states $|\psi_1(t)\rangle$ and $|\psi_3(t)\rangle$ correspond to spin-up (with $+\hbar/2$ values) while $|\psi_2(t)\rangle$ and $|\psi_4(t)\rangle$ correspond to spin-down (with $-\hbar/2$ values). In fact, and as will be discussed below (see Section 12.4.13), the Dirac equation does describe the motion of electrons and that the spin degree of freedom comes out naturally from the formalism, which was not the case with the Schrödinger equation where, as shown in Chapter 6, the spin had to be added by hand resulting in the Pauli equation (see Section 6.4 on Page 403).

Second, and as we shall see in Subsections 12.4.10–12.4.12, the negative energy solutions correspond to physical solutions that describe *antiparticles*. As seen in Chapter 6, when applied to an electron where the spin is taken into account, the Schrödinger equation yields a wave function that has two components corresponding to the two spin orientations (spin-up and spin-down). However, in the relativistic description of a Dirac particle, the Dirac equation yields a 4-component spinor; the two extra degrees of freedom in the Dirac 4-spinor are due to the existence of a positron, the antiparticle of the electron; these two additional degrees of freedom account for the spin-up and spin-down orientations of the positron's spin. One of the historical hallmarks of the Dirac equation is its overarching predictive power about the existence of antiparticles which come out naturally from the formalism.

In Subsections 12.4.10–12.4.12, we will discuss in more detail the conceptual challenges encountered in interpreting the negative energy solutions. We will see that a proper resolution of these challenges warrants going beyond the single particle interpretation of the Dirac equation.

Finally, we can summarize the solutions of the Dirac equation, when applied to a free particle at rest, as follows:

- The respective states $|\psi_1\rangle$ and $|\psi_2\rangle$ listed in (12.297) describe the spin-up and spin-down orientations of a spin-1/2 *particle* with a positive energy $E^{(+)} = mc^2$.
- The respective states $|\psi_3\rangle$ and $|\psi_4\rangle$ listed in (12.298) describe the spin-up and spin-down orientations of a spin-1/2 *antiparticle* with negative energy $E^{(-)} = -mc^2$.

Hence, the Dirac equation describes spin 1/2 particles and spin 1/2 antiparticles. For each energy sign, there are two linearly independent solutions corresponding to the two spin states (up and down) as expected for a spin 1/2 particle such as an electron or an antiparticle such as a positron. After all, Dirac's aim from the very outset was to come up with a relativistic quantum theory of electrons.

It helps to notice that the spinors $|u_1\rangle$, $|u_2\rangle$, $|u_3\rangle$, $|u_4\rangle$ are orthonormal

$$\langle u_\alpha \mid u_\beta \rangle = \delta\alpha\beta, \qquad \alpha, \beta = 1, 2, 3, 4, \tag{12.299}$$

and that they form a complete set

$$\sum_{\alpha=1}^{4} |u_\alpha\rangle\langle u_\alpha| = I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{12.300}$$

This shows that we cannot ignore the negative energy states since the positive energy solutions alone (i.e., $|u_1\rangle$ and $|u_2\rangle$) do not form a complete set of solutions; however, the four states $|u_1\rangle$, $|u_2\rangle$, $|u_3\rangle$, $|u_4\rangle$ do form a complete set of states. Thus, the spinor of any free particle at rest fcan be ex-pressed as a linear combination of the states (12.297)–(12.298):

$$|\psi(t)\rangle = a_1|\psi_1(t)\rangle + a_2|\psi_2(t)\rangle + a_3|\psi_3(t)\rangle + a_4|\psi_4(t)\rangle, \tag{12.301}$$

where the expansion coefficients a_1, a_2, a_3, a_4 are complex numbers.

Example 12.6

Use Heisenberg's matrix formulation of quantum mechanics to find the energy levels and eigenstates of a particle at rest.

Solution

The Heisenberg approach consists in solving the eigenvalue problem $\hat{H}|\psi\rangle = E|\psi\rangle$ by diagonalizing the Hamiltonian's matrix. For the case of a particle at rest, inserting $\hat{\vec{p}} = 0$ into Dirac's Hamilonian (12.147)

$$\hat{H} = c\vec{\alpha} \cdot \hat{\vec{p}} + mc^2\beta = mc^2\beta, \tag{12.302}$$

and using the matrix form (12.161) of β, we can write the Hamiltonian in a matrix form as follows

$$\hat{H} = mc^2\beta = mc^2\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{12.303}$$

Diagonalizing this simple matrix, we obtain two doubly degenerate solutions with positive and negative energies, $E_1 = E_2 \equiv E^{(+)} = mc^2$ and $E_3 = E_4 \equiv E^{(-)} = -mc^2$, respectively; that is, we obtain the following eigenvalues along with their respective (normalized) eigenstates

$$E^{(+)} = mc^2, \qquad |u_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad |u_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tag{12.304}$$

$$E^{(-)} = -mc^2, \qquad |u_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |u_4\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{12.305}$$

To obtain the time-dependent solutions, we simply need to multiply the above eigenstates by the corresponding exponents; this leads to equations (12.297)–(12.298).

12.4.9 Solutions of the Dirac Equation for a Free Particle in Motion

Let us now consider the general case of a free particle with non-zero momentum, $\vec{p} \neq 0$ where $\vec{p} = p_x\hat{i} + p_y\hat{j} + p_z\hat{k}$; since the particle is free, its momentum's magnitude is constant, p = contant. The solutions of the Dirac equation (12.232) for a free particle must be plane waves[33] that are eigenfunctions to the momentum operator $\hat{p}_\mu$:

$$\psi(x^\mu) = u(p^\mu)e^{-ip_\mu x^\mu/\hbar} = u(p^\mu)e^{-i(Et-\vec{p}\cdot\vec{r})/\hbar}, \tag{12.306}$$

where the amplitude $u(p^\mu) \equiv u(E, \vec{p})$ is a 4-component spinor where its components depend on the 4-momentum p^μ; they depend neither on $\vec{r}$ nor on t. For the sake of notational brevity, and whenever warranted, we will use $\psi(x)$ and $u(p)$ to denote $\psi(x^\mu)$ and $u(p^\mu)$, respectively. The spinor $u(p^\mu)$ is represented (in the momentum representation) by a four-component column matrix:

$$u(p^\mu) = \langle p^\mu|u\rangle = \begin{pmatrix} u_1(p^\mu) \\ u_2(p^\mu)) \\ u_3(p^\mu) \\ u_4(p^\mu) \end{pmatrix}, \tag{12.307}$$

where $u_j(p^\mu)$ (with $j = 1, 2, 3, 4$) are scalar (constant) functions which depend on the particle's energy E and its 3-momentum $\vec{p}$; i.e., $u_j(p^\mu) \equiv u_j(E, \vec{p})$.

[33] Recalling that, since $p_\mu = (E/c, -\vec{p})$ and $x^\mu = (ct, \vec{r})$, we have: $-ip_\mu x^\mu/\hbar = -i(Et - \vec{p}\cdot\vec{r})/\hbar$, which represents the phase of the oscillating plane wave.

As mentioned above, the state $\psi(x^\mu)$ displayed in equation (12.306) is an eigenfunction of the momentum operator $\hat{p}^\mu$, since

$$\begin{aligned}\hat{p}^\mu\psi(x^\mu) &= i\hbar\partial^\mu\psi(x^\mu) = \left(\frac{i\hbar}{c}\frac{\partial\psi(x^\mu)}{\partial t}, -i\hbar\vec{\nabla}\psi(x^\mu)\right)\\ &= u(p^\mu)\left(\frac{i\hbar}{c}\frac{\partial e^{-i(Et-\vec{p}\cdot\vec{r})/\hbar}}{\partial t}, -i\hbar\vec{\nabla}e^{-i(Et-\vec{p}\cdot\vec{r})/\hbar}\right)\\ &= \left(\frac{E}{c}, \vec{p}\right)u(p^\mu)e^{-i(Et-\vec{p}\cdot\vec{r})/\hbar}\\ &= p^\mu\psi(x^\mu),\end{aligned} \tag{12.308}$$

where $p^\mu = \left(\frac{E}{c}, \vec{p}\right)$ is just the 4-momentum.

Substituting (12.306) into (12.231), we obtain

$$\left(i\hbar\gamma^\mu\partial_\mu - mc\right)\psi(x^\mu) = u(p^\mu)\left(i\hbar\gamma^\mu\partial_\mu - mc\right)e^{-i(Et-\vec{p}\cdot\vec{r})/\hbar} = 0. \tag{12.309}$$

Using the fact that (see (12.197)):

$$\gamma^\mu\partial_\mu = \frac{1}{c}\gamma^0\frac{\partial}{\partial t} + \vec{\gamma}\cdot\vec{\nabla} = \frac{1}{c}\gamma^0\frac{\partial}{\partial t} + \gamma^1\frac{\partial}{\partial x} + \gamma^2\frac{\partial}{\partial y} + \gamma^3\frac{\partial}{\partial z}, \tag{12.310}$$

we have

$$\begin{aligned}i\hbar\gamma^\mu\partial_\mu\psi(x^\mu) &= i\hbar u(p^\mu)\left(\frac{1}{c}\gamma^0\frac{\partial}{\partial t} + \gamma^1\frac{\partial}{\partial x} + \gamma^2\frac{\partial}{\partial y} + \gamma^3\frac{\partial}{\partial z}\right)e^{-i(Et-\vec{p}\cdot\vec{r})/\hbar}\\ &= u(p^\mu)\left(\frac{E}{c}\gamma^0 - \gamma^1p_x - \gamma^2p_y - \gamma^3p_z\right)e^{-i(Et-\vec{p}\cdot\vec{r})/\hbar}.\end{aligned} \tag{12.311}$$

Notice that the plane wave solutions (12.306) are joint eigenfunctions of the energy operator, $i\hbar\partial/\partial t$, and momentum operator, $-i\hbar\vec{\nabla}$, with eigenvalues E and $\vec{p}$, respectively. Inserting (12.311) into (12.309) and canceling the exponent, we obtain

$$\left(\frac{E}{c}\gamma^0 - \gamma^1p_x - \gamma^2p_y - \gamma^3p_z - mc\right)u(p^\mu) = 0; \tag{12.312}$$

this reduces to

$$\left(\gamma^0E - mc^2 - c\gamma^1p_x - c\gamma^2p_y - c\gamma^3p_z\right)u(p^\mu) = 0. \tag{12.313}$$

Now, substituting the matrix forms (12.222)–(12.223) of γ^0, γ^1, γ^2, γ^3 and the matrix form (12.307) of $u(p^\mu)$ into (12.313), we obtain

$$\begin{pmatrix} E-mc^2 & 0 & -cp_z & -c(p_x-ip_y)\\ 0 & E-mc^2 & -c(p_x+ip_y) & cp_z\\ cp_z & c(p_x-ip_y) & -(E+mc^2) & 0\\ c(p_x+ip_y) & -cp_z & 0 & -(E+mc^2)\end{pmatrix}\begin{pmatrix}u_1\\u_2\\u_3\\u_4\end{pmatrix} = 0, \tag{12.314}$$

which reduces to four coupled, homogeneous equations for the four components of u:

$$(E-mc^2)u_1 - cp_zu_3 - c(p_x-ip_y)u_4 = 0, \tag{12.315}$$

$$(E-mc^2)u_2 - c(p_x+ip_y)u_3 + cp_zu_4 = 0, \tag{12.316}$$

$$cp_zu_1 + c(p_x-ip_y)u_2 - (E+mc^2)u_3 = 0, \tag{12.317}$$

$$c(p_x+ip_y)u_1 - cp_zu_2 - (E+mc^2)u_4 = 0. \tag{12.318}$$

For this system of four linear homogeneous equations to have non-trivial solutions, its determinant must be zero:

$$\left[E^2 - m^2c^4 - \left(c^2p_x^2 + c^2p_y^2 + c^2p_z^2\right)\right]^2 = \left(E^2 - m^2c^4 - c^2\vec{p}^{\,2}\right)^2 = 0, \tag{12.319}$$

which is a fourth-order equation in p. As expected, this yields two doubly degenerate energy values, very much as in the case of the particle at rest:

$$E_1 = E_2 \equiv E^{(+)} = \sqrt{c^2\vec{p}^{\,2} + m^2c^4}, \qquad E_3 = E_4 \equiv E^{(-)} = -\sqrt{c^2\vec{p}^{\,2} + m^2c^4}. \tag{12.320}$$

Hence, the Dirac solutions for a free particle consists of four eigenstates, two linearly independent solutions with positive energy $E^{(+)}$ and two other linearly independent solutions with negative energy $E^{(-)}$, very much like the case discussed above for a particle at rest.

We now need to find the eigenstates of (12.315)–(12.318). This task requires some involved, yet straightforward, calculations; we have carried them out in Problem (12.12) on Page 833 where we have shown that the solutions of (12.315)–(12.318) consist of four linearly independent eigenstates:

- two degenerate *positive energy* states $|u_1\rangle$ and $|u_2\rangle$ describing, respectively, spin-up and spin-down orientations of a spin-1/2 *particle*,

$$\boxed{u_1(p^\mu) = N\begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \end{pmatrix}, \quad u_2(p^\mu) = N\begin{pmatrix} 0 \\ 1 \\ \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix}, \quad E^{(+)} = \sqrt{c^2\vec{p}^{\,2} + m^2c^4},} \tag{12.321}$$

- and two degenerate *negative energy* states $|u_3\rangle$ and $|u_4\rangle$ describing, respectively, spin-up and spin-down orientations of a spin-1/2 *antiparticle*,

$$\boxed{u_3(p^\mu) = N\begin{pmatrix} \frac{cp_z}{E^{(-)}-mc^2} \\ \frac{c(p_x+ip_y)}{E^{(-)}-mc^2} \\ 1 \\ 0 \end{pmatrix}, \quad u_4(p^\mu) = N\begin{pmatrix} \frac{c(p_x-ip_y)}{E^{(-)}-mc^2} \\ \frac{-cp_z}{E^{(-)}-mc^2} \\ 0 \\ 1 \end{pmatrix}, \quad E^{(-)} = -\sqrt{c^2\vec{p}^{\,2} + m^2c^4}.} \tag{12.322}$$

The normalization constant N is given by equation (12.722): $N = \sqrt{E^{(+)} + mc^2}$. Note that if we plug $p_x = p_y = p_z = 0$ in the four equations above, we recover the eigenstates (12.296) for a particle at rest where $|u_1\rangle$ and $|u_2\rangle$ correspond to the positive energy solutions while $|u_3\rangle$ and $|u_4\rangle$ to the negative energy solutions.

Finally, inserting (12.321) and (12.322) into (12.306), we obtain the time-dependent solutions of the Dirac equation for a free particle with $\vec{p} \neq 0$; they consist of

- two degenerate *positive energy* solutions describing, respectively, spin-up and spin-down orientations of a spin-1/2 *particle*,

$$\boxed{\psi_1(x^\mu) = u_1(p^\mu)\, e^{-i\left(E^{(+)}t - \vec{p}\cdot\vec{r}\right)/\hbar}, \qquad \psi_2(x^\mu) = u_2(p^\mu)\, e^{-i\left(E^{(+)}t - \vec{p}\cdot\vec{r}\right)/\hbar},} \tag{12.323}$$

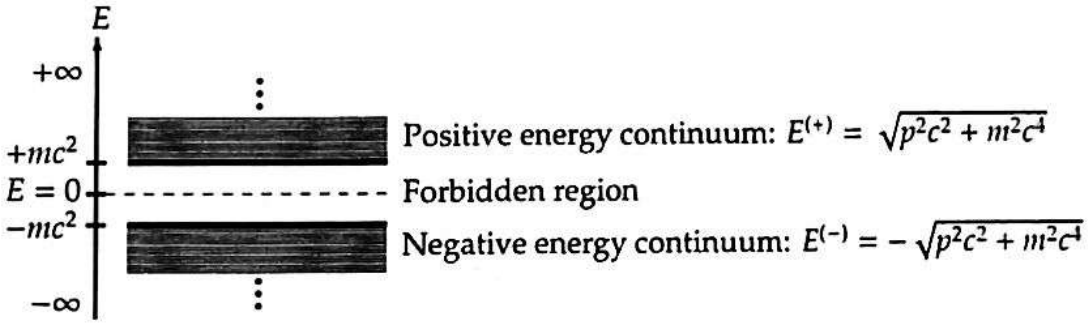


Figure 12.2: Energy spectrum of a free Dirac particle: It consists of two energy continua (allowed regions) that are separated by a forbidden region of width $2mc^2$; the positive energy band, $E^{(+)} > mc^2$, corresponds to particles and the negative energy band, $E^{(-)} < -mc^2$, to antiparticles.

- and two degenerate *negative energy* solutions describing, respectively, spin-up and spin-down orientations of a spin-1/2 *antiparticle*,

$$\boxed{\psi_3(x^\mu) = u_3(p^\mu)\, e^{-i(E^{(-)}t - \vec{p}\cdot\vec{r})/\hbar}, \qquad \psi_4(x^\mu) = u_4(p^\mu)\, e^{-i(E^{(-)}t - \vec{p}\cdot\vec{r})/\hbar}.} \tag{12.324}$$

As in the case of the Klein–Gordon equation, the Dirac equation also yields negative energy solutions; we cannot ignore them since the positive energy solutions do not form a complete set. In the following section, we are going to discuss in more detail the interpretation of the negative energy solutions.

12.4.10 Negative-Energies Interpretation: Dirac's Hole Theory

The energy spectrum of a free Dirac particle (see Eq. (12.320)) is continuous, ranging from $-\infty$ to $+\infty$, except for the forbidden energy gap located between $-mc^2 < E < mc^2$ as shown in Figure 12.2. So we have three regions: two allowed regions (one with a positive energy continuum $E > mc^2$ and another with a negative energy continuum $E < -mc^2$) and one forbidden region (with energies between $-mc^2 < E < mc^2$). The existence of this forbidden energy gap is due to the fact that, since the energies of the free particle are given by $E^{(\pm)} = \pm\sqrt{c^2p^2 + m^2c^4}$, the (continuous) negative energy spectrum, although not bounded below the $-mc^2$ level, is capped at the top by $E^{(-)}_{max} = -mc^2$, and the positive energy spectrum is bounded at the bottom by $E^{(+)}_{min} = mc^2$. For instance, in the case of a free electron, this energy gap is

$$2m_ec^2 \simeq 2 \times 0.511\ \text{MeV} = 1.022\ \text{MeV}. \tag{12.325}$$

The real issue now is how to interpret the negative energy solutions of the Dirac equation. In the literature, one can find several approaches that have provided interpretations for the Dirac's negative energies. In what follows, we are going to consider the three most known interpretational approaches: (a) the *Dirac hole theory*, (b) the *Feynman-Stueckelber interpretentation*, and (c) the *Klein paradox*.

Dirac's Hole Theory: Interpretation of the Negative Energies

It is instructive to find out, at least from a historical perspective, how has Dirac dealt with the negative energies. To interpret them, Dirac has formulated his *hole theory*[34] in which he has postulated that the vacuum is composed of an *infinite continuum of negative energy levels* and that each level is filled with two electrons of opposite spins so as to be compliant with the Pauli exclusion principle. This infinite continuum is called the *Dirac sea*[35] or the Dirac vacuum. While all the negative energy levels are fully occupied by electrons, the positive energy levels are empty as illustrated in Figure 12.3-(a).

Since no two electrons can occupy the same state[36] in the Dirac sea, as dictated by the Pauli principle, any additional electron must occupy a positive energy state, for it cannot fall into one of the negative energy levels. If we supply enough energy, a minimum[37] of $2mc^2$ (supplied, for instance, by a γ-ray photon), a negative energy electron can absorb this photon and make a transition to a positive energy state leaving behind a "hole" in the negative energy sea; measured with respect to the vacuum, this hole has positive charge and positive energy. In this process, an electron-hole pair is created where we end up with a positive energy electron and a hole in the negative energy sea as shown in Figure 12.3-(b); the hole in the sea appears to have features and properties that are opposite to the Dirac's sea electrons (i.e., positive charge, positive energy, and negative momentum). Dirac interpreted the hole as a positive energy "particle" having the same mass as the electron but with an opposite (or positive) charge. This idea has led Dirac to predict the **positron**[38]. Dirac's 1930 prediction on the existence of the *antielectron, or the antimatter* counterpart of the electron, was confirmed experimentally by Anderson[39] in 1933 when he discovered the positron in cosmic rays. Throughout this disquisition, we will be denoting an electron and a positron by e^- and e^+, respectively. The prediction of the positron by Dirac and its discovery by Anderson have given considerable support to the Dirac equation.

The above discussed process in which a radiation, such as a γ-ray, of energy greater than $2mc^2$ is absorbed by a negative energy electron and resulting, as illustrated in Figure 12.3-(b), in the creation of a positive energy electron and a positive energy positron (i.e., the negative energy hole) is called *pair creation*. We should note that, based on conservation laws, a γ-ray cannot vanish on its own in free space and creating an electron-positron pair, since the energy and momentum cannot be jointly conserved in vacuum; an external field, such as that of an atomic nucleus, is required to absorb some of the photon's momentum:

$$\text{Pair Creation:} \qquad \gamma + \text{Nucleus} \longrightarrow e^- + e^+ + \text{Nucleus}^*, \tag{12.326}$$

where Nucleus* denotes the original nucleus after having acquired some momentum from the photon. Pair-creation can only occur in the presence of an external field making it possible for the momentum and energy to be both conserved. The reverse process is called *pair annihilation*:

---

[34] P. A. M. Dirac, A theory of electrons and protons, *Proc. R. Soc.* (London) **A126**, 360–365 (1930).

[35] While having an infinite negative charge and an infinite negative energy, the Dirac sea is not physically observable, for it produces no measurable field. First, since the infinite sea of negative-energy electrons produces a uniform charge density everywhere in space, the electric field it creates at the center of a "fictitious" uniformly charged sphere of infinite extent is zero as per Gauss's law. Second, the energy of the Dirac sea is not a physically observable quantity, since one cannot measure absolute energies, one can only measure energy differences.

[36] Recall that each level consists of two states (spin up and spin down) and that each state is occupied by at most one electron.

[37] In the case of an electron, this energy is about $2mc^2 \simeq 1$ MeV which can be supplied by energetic γ-ray photons.

[38] P. A. M. Dirac, *Proc. Roy. Soc.* (London) **A 126**, 360 (1930); **A 126**, 801 (1930); **A 133**, 821 (1931).

[39] Carl D. Anderson, *Phys. Rev.*, **43**, 491 (1933).

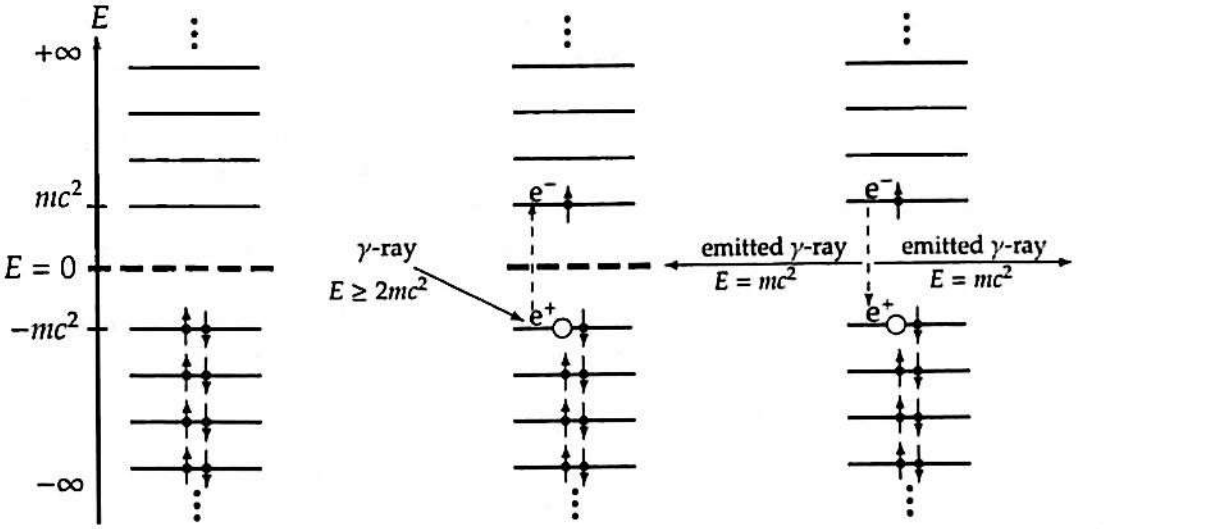


Figure 12.3: **(a) Dirac Sea (vacuum):** consists of an infinite continuum of negative energy states $-\infty < E \leq -mc^2$ that are fully occupied by electrons and the positive energy states empty. **(b) Electron-Positron Pair Creation:** A γ-ray (with a minimum energy of $2mc^2$) is absorbed by a negative energy electron and making a transition to a positive energy state leaving behind a positron (positively charged hole). **(c) Electron-Positron Pair Annihilation:** A positive energy electron annihilates a positron by emitting two γ-rays, each with energy mc^2, moving in opposite directions to conserve momentum.

When a positive energy electron meets a positron (i.e., when a positive energy electron falls into a hole), both of them annihilate by emitting a radiation with a minimum energy of $2mc^2$; this radiation consists of two[40] γ-ray photons each with energy mc^2 as illustrated in Figure 12.3-(c):

$$\text{Pair Annihilation:} \qquad e^- + e^+ \longrightarrow \gamma + \gamma. \tag{12.327}$$

Note that, since two photons are emitted during the electron-positron annihilation process, both the energy and momentum can be jointly conserved without the need for an external field to absorb some of the momentum.

It is worth mentioning that pair production and pair annihilation processes have become so common in today's experiments that we take them for granted. However, for these processes to have been correctly predicted almost a century ago by Dirac's hole theory is a testimony of human genius[41]. The Dirac equation turned out to be not a mere marriage between special relativity and quantum mechanics, it also introduced the new concept about the non-conservation of the number of particles. By predicting the existence of antiparticles and pair-production, the Dirac equation has offered a remarkable paradigm shift from the traditional theory of non-relativistic quantum mechanics.

However, in spite of its initial technical success in interpreting the negative energy solutions of spin-1/2 particles (fermions), Dirac's hole theory suffers from a number of problems.

[40] With only one photon emitted, only charge and energy can be conserved; but to conserve momentum, two photons moving in opposite directions must be emitted.

[41] On a lighter note, in one of his talks, Dirac remarked that his equation was more intelligent than its author.

First, it does not work for bosons since bosons do not obey the Pauli exclusion principle (only fermions do). Second, even though the infinite negative charge and negative energy of the Dirac sea are not physically observable quantities as discussed above, the very idea of an infinite sea of negative energy states filled with electrons raises a number of conceptual problems. For instance, the hole theory provides no explanation about the interactions between the electrons occupying the negative energy states in the sea. Third, although the Dirac equation started as a one particle theory, we ended up with a theory that is multiparticle in nature which involves an infinite sea of electrons and processes such as creation and annihilation of particles. So, in spite of its technical significance, the Dirac hole theory, which rests on the idea of an infinite sea of negative energy states that are occupied by electrons, is unsatisfactory; *its significance is merely historical.*

In the next Subsection we are going to deal with the Feynman-Stückelberg approach that applies to both bosons and fermions and which offers a more general interpretation of the negative energy states with no need for an infinite sea of negative energy states; this method treats electrons and positrons on equal footing with no need for the concept of holes.

12.4.11 Negative-Energies Interpretation: Feynman-Stückelberg Method

This prescription, which was proposed by Stüeckelberg (1941) and by Feynman (1949), rests on causality, not on the Pauli exclusion principle nor on Dirac's concept of an infinite sea of electrons. The Feynman-Stückelberg Interpretation resolves the conceptual difficulties inherent in the Dirac's hole theory by going beyond the single particle interpretation of the Dirac equation; it interprets the negative energy states by introducing the concept of antiparticles which, inevitably, leads to the non-conservation of the particle number. In what follows, we are going to briefly examine the works of Stückelberg and Feynman and then draw conclusions about their common results.

First, instead of dealing with single particle states, Ernst Stüeckelberg[42] used fields. To ensure consistency with causality, which stipulates that positive energy states propagate forwards in time, Stüeckelberg interpreted the negative energy solutions as corresponding to negative energy modes of the electron field propagating backward in time. Theoretically, the time dependence, $e^{-iE^{(+)}t}$ (where $t > 0$), of a positive energy state that is propagating forwards in time is mathematically identical to that of a negative energy state that is propagating backwards in time as depicted pictorially in Figure 12.4-(a); this idea can be implemented by simply reversing the signs of the energy and time, simultaneously: $t \longrightarrow -t$ and $E^{(+)} \longrightarrow -E^{(+)}$ (or $E^{(+)} \longrightarrow E^{(-)}$); that is

$$e^{-i(-E^{(+)})(-t)} \equiv e^{-iE^{(+)}t}, \qquad \text{where} \quad t > 0. \tag{12.328}$$

This result is consistent with the *causality requirement where positive energy states only propagate forwards in time and negative energy states travel only backwards in time.* So, according to the Stüeckelberg interpretation, *the negative energy solutions can describe either a negative energy particle (i.e., electron) propagating backwards in time, or a positive energy antiparticle (positron) propagating forward in time.* In this way, he offered a satisfactory pictorial view of particles and antiparticles by arguing that a particle and its antiparticle have equal mass, equal spin, but opposite charges. In his pictorial interpretation, Stüeckelberg has constructed diagrams that

[42] Ernst Stüeckelberg, *Helvetica Physica Acta*, **Vol.14**, 51 (1941).

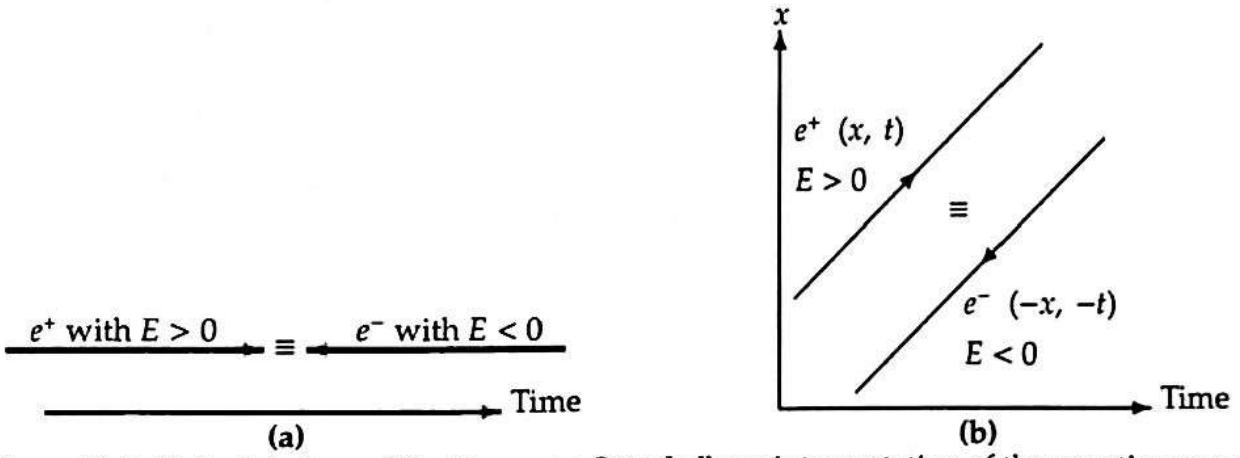


Figure 12.4: Pictorial view of the Feynman-Stüeckelberg interpretation of the negative energy solutions: **(a):** using only the time variable, **(b):** using space and time: The solution of a negative energy particle (electron) traveling backwards in space and time $(-x, -t)$ is equivalent to that of a positive energy antiparticle (positron) traveling forwards in space and time (x, t).

allowed him to perform perturbation theory calculations to extract quantitative information about particle-antiparticle interactions and scattering processes.

Second, it was Feynman[43] who has managed to systematize perturbation theory calculations by introducing his famous diagrams – the Feynman diagrams[44] – where each diagram's line represents a particle that propagates either forward or backward in time as shown in Figure 12.4. In his 1949 theory of positrons, Feynman proposed the idea that the negative energy states may be pictured in spacetime, as suggested by Stückelberg in 1941, as waves traveling backwards in time and away from the external potential; he viewed this traveling wave as corresponding to a positron approaching the external potential and annihilating the electron.

To summarize, the Feynman-Stüeckelberg interpretation is founded on the idea that *a negative energy solution represents a negative energy particle propagating backwards in space and time or, equivalently, a positive energy antiparticle traveling forwards in space and time*; Figure 12.4-(b) provides a pictorial representation, since the space and time dependence of a positive energy solution of an antiparticle that propagates forwards in space and time is mathematically identical to that of the negative energy solution of its particle that propagates backwards in space and time

$$\exp\left[-i\left(E^{(+)}t - \vec{p}\cdot\vec{r}\right)\right] \equiv \exp\left[-i\left((-E^{(+)})(-t) - (-\vec{p})\cdot(-\vec{r})\right)\right], \qquad \text{where} \qquad t > 0; \quad (12.329)$$

the left-hand side term $\exp\left[-i\left(E^{(+)}t - \vec{p}\cdot\vec{r}\right)\right]$ describes a propagation that is forward in space and in time, and the right-hand side term $\exp\left[-i\left((-E^{(+)})(-t) - (-\vec{p})\cdot(-\vec{r})\right)\right]$ describes a propagation that is backward in space and in time.

Let us now apply this idea to the solutions of a free Dirac particle obtained in Subsection 12.4.9. The Feynman-Stüeckelberg interpretation considers the two negative-energy

[43] Richard Feynman, *Physical Review*, **76**, 749 (1949).

[44] On a side note, Feynman diagrams turned out to be a very powerful tool that greatly simplified calculations in quantum field theory.

antiparticle states $|\psi_3\rangle$ and $|\psi_4\rangle$ as having *negative time evolution* and rewriting them in a way to represent positive energy antiparticles. To implement this idea, we define antipaticle states by simply *reversing the signs of all four components of the 4-momentum (i.e., the energy and the 3-momentum) and flipping the spin direction in* Eqs. (12.322) and (12.324). That is, instead of $\psi_3(x^\mu)$ and $\psi_4(x^\mu)$, we can construct new antiparticle plane wave solutions $\Phi_3(x^\mu)$ and $\Phi_4(x^\mu)$ by introducing two new amplitudes $v_1(p^\mu)$ and $v_2(p^\mu)$ and expressing them in terms of $u_4(-p^\mu)$ and $u_3(-p^\mu)$ as follows: $v_1(p^\mu) \equiv u_4(-p^\mu)$ and $v_2(p^\mu) \equiv u_3(-p^\mu)$; that is,

$$\Phi_3(x^\mu) = v_1(p_\mu)\, e^{ip_\mu x^\mu} \equiv u_4(-p_\mu)\, e^{-i(-p_\mu)x^\mu} = u_4(-p_\mu)\, e^{ip_\mu x^\mu}, \tag{12.330}$$

$$\Phi_4(x^\mu) = v_2(p_\mu)\, e^{ip_\mu x^\mu} \equiv u_3(-p_\mu)\, e^{-i(-p_\mu)x^\mu} = u_3(-p_\mu)\, e^{ip_\mu x^\mu}, \tag{12.331}$$

or

$$\Phi_3(x^\mu) = v_1(E^{(+)}, \vec{p})\, e^{i(E^{(+)}t - \vec{p}\cdot\vec{r})/\hbar} \equiv u_4(-E^{(+)}, -\vec{p})\, e^{-i\left[(-E^{(+)})t - (-\vec{p})\cdot\vec{r}\right]/\hbar} = u_4(-E^{(+)}, -\vec{p})\, e^{i(E^{(+)}t - \vec{p}\cdot\vec{r})/\hbar}, \tag{12.332}$$

$$\Phi_4(x^\mu) = v_2(E^{(+)}, \vec{p})\, e^{i(E^{(+)}t - \vec{p}\cdot\vec{r})/\hbar} \equiv u_3(-E^{(+)}, -\vec{p})\, e^{-i\left[(-E^{(+)})t - (-\vec{p})\cdot\vec{r}\right]/\hbar} = u_3(-E^{(+)}, -\vec{p})\, e^{i(E^{(+)}t - \vec{p}\cdot\vec{r})/\hbar}, \tag{12.333}$$

where $v_1(p^\mu)$ and $v_2(p^\mu)$ can be obtained explicitly from (12.322):

$$v_1(p^\mu) \equiv u_4(-E^{(+)}, -\vec{p}) = N \begin{pmatrix} \frac{c[(-p_x)-i(-p_y)]}{-E^{(+)}-mc^2} \\ \frac{-c(-p_z)}{-E^{(+)}-mc^2} \\ 0 \\ 1 \end{pmatrix} = N \begin{pmatrix} \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix}, \tag{12.334}$$

$$v_2(p^\mu) \equiv u_3(-E^{(+)}, -\vec{p}) = N \begin{pmatrix} \frac{c(-p_z)}{-E^{(+)}-mc^2} \\ \frac{c[(-p_x)+i(-p_y)]}{-E^{(+)}-mc^2} \\ 1 \\ 0 \end{pmatrix} = N \begin{pmatrix} \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \\ 1 \\ 0 \end{pmatrix}, \tag{12.335}$$

where $E^{(+)} = \sqrt{c^2\vec{p}^{\,2} + m^2c^4}$ and the normalization constant N is given by $N = \sqrt{E^{(+)} + mc^2}$ (see equation (12.722) for the derivation of N).

So, the Dirac equation for a free particle with $\vec{p} \neq 0$ has four linearly independent solutions that can be selected as follows:

- we can select either the set $u_1(p^\mu)$, $u_2(p^\mu)$, $u_3(p^\mu)$, $u_4(p^\mu)$ (listed in equations (12.321)-(12.322)) where u_1 and u_2 respectively describe spin-up and spin-down *particles* with *positive energy* $E^{(+)}$, while u_3 and u_4 respectively describe spin-up and spin-down *antiparticles* with *negative energy* $E^{(-)}$,

- or we can select the set $u_1(p^\mu)$, $u_2(p^\mu)$, $v_1(p^\mu)$, $v_2(p^\mu)$ where u_1 and u_2 respectively describe spin-up and spin-down *particles* with *positive energy* $E^{(+)}$, while v_1 and v_2 respectively describe spin-up and spin-down *antiparticles* with *positive energy*, $E^{(+)}$, *traveling forwards*

in time:

$$u_1 = N\begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \end{pmatrix}, \quad u_2 = N\begin{pmatrix} 0 \\ 1 \\ \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix}, \quad v_1 = N\begin{pmatrix} \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = N\begin{pmatrix} \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \\ 1 \\ 0 \end{pmatrix}. \tag{12.336}$$

Clearly, the set $u_1(p^\mu)$, $u_2(p^\mu)$, $v_1(p^\mu)$, $v_2(p^\mu)$ is the most appropriate since it involves only positive energy solutions.

Summary of the Dirac Solutions for a Free Particle:
Finally, to obtain the plane wave solutions for a free particle, we need simply to insert (12.321) into (12.323), (12.334) into (12.332), and (12.335) into (12.333); these yield, respectively:

- two degenerate positive energy solutions describing, respectively, spin-up and spin-down orientations of an *electron (particle)*

$$\psi_1(x^\mu) = N\begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \end{pmatrix} e^{-i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}, \qquad \psi_2(x^\mu) = N\begin{pmatrix} 0 \\ 1 \\ \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix} e^{-i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}, \tag{12.337}$$

where, as Eq (12.344) shows, u_1 and u_2 satisfy: $\boxed{\left(\gamma^\mu p_\mu - mc\right)u_j = 0}$ with $j = 1, 2$,

- and two degenerate positive energy solutions describing, respectively, spin-up and spin-down orientations of a *positron (antiparticle)*

$$\Phi_3(x^\mu) = N\begin{pmatrix} \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix} e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}, \qquad \Phi_4(x^\mu) = N\begin{pmatrix} \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \\ 1 \\ 0 \end{pmatrix} e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}, \tag{12.338}$$

where, as Eq (12.345) shows, v_1 and v_2 satisfy: $\boxed{\left(\gamma^\mu p_\mu + mc\right)v_j = 0}$ with $j = 1, 2$.

Remarks

- As shown in Example 12.7, we can verify that both of the *free particle* states (12.337) and the *antiparticle* states (12.338) correspond to *positive energy*, $E^{(+)}$, and *positive momentum*, $\vec{p}$, solutions, since

$$\hat{H}\psi_j(x^\mu) = E^{(+)}\psi_j(x^\mu), \qquad \hat{\vec{p}}\,\psi_j(x^\mu) = \vec{p}\,\psi_j(x^\mu), \qquad j = 1, 2. \tag{12.339}$$

$$\hat{H}^{(A)}\Phi_j(x^\mu) = E^{(+)}\Phi_j(x^\mu), \qquad \hat{\vec{p}}^{\,(A)}\,\Phi_j(x^\mu) = \vec{p}\,\Phi_j(x^\mu), \qquad j = 3, 4. \tag{12.340}$$

Notice that, to be in conformity with the Feynman-Stüeckelberg interpretation where the signs of the energy and momentum are reversed for antiparticles (i.e., $(E, \vec{p}) \longrightarrow (-E, -\vec{p})$), we have *redefined the antiparticle energy and momentum operators* in (12.340) as follows :

$$\hat{H}^{(A)} = -i\hbar\frac{\partial}{\partial t}, \qquad \hat{\vec{p}}^{(A)} = i\hbar\vec{\nabla}. \tag{12.341}$$

In this way, we obtain physically meaningful expressions for the antiparticle's energy and momentum which must be both positive as stipulated by the Feynman-Stüeckelberg interpretation. Had we taken the normal quantum mechanical definitions of the energy and momentum operators (i.e., $\hat{H} = i\hbar\partial/\partial t$ and $\hat{\vec{p}} = -i\hbar\vec{\nabla}$), we would have ended up with negative energy and negative momentum expressions

$$\hat{H}\Phi_j(x^\mu) = -E^{(+)}\Phi_j(x^\mu), \qquad \hat{\vec{p}}\,\Phi_j(x^\mu) = -\vec{p}\,\Phi_j(x^\mu), \qquad j = 3,\ 4, \tag{12.342}$$

which are unphysical.

- By analogy to (12.341), using the transformation $(E, \vec{p}) \longrightarrow (-E, -\vec{p})$, we can show that the orbital angular momentum operator $\hat{\vec{L}}^{(A)}$ for antiparticles is equal to $-\hat{\vec{L}}$ for particles:

$$\hat{\vec{L}} = \vec{r}\times\vec{p} \qquad \longrightarrow \qquad \hat{\vec{L}}^{(A)} = \vec{r}\times(-\vec{p}) = -\vec{r}\times\vec{p} \equiv -\hat{\vec{L}}. \tag{12.343}$$

Similarly, we will show below (see equation (12.394)) that the spin operator $\hat{\vec{S}}^{(A)}$ for antiparticles is equal to $-\hat{\vec{S}}$ for particles; that is, the spin flips for antiparticles.

- We should also note that, if we substitute (12.337) into the Dirac equation (12.231), we can verify that the *particle spinors* u_1 and u_2 satisfy the following equation

$$\left(\gamma^\mu p_\mu - mc\right)u_j(p_\mu) = 0, \qquad j = 1,\ 2, \tag{12.344}$$

and if we substitute (12.338) into (12.231), and as shown in Example 12.8, we see that the *antiparticle spinors* v_1 and v_2 obey this equation

$$\left(\gamma^\mu p_\mu + mc\right)v_j(p_\mu) = 0, \qquad j = 1,\ 2. \tag{12.345}$$

Alternatively, we can obtain (12.345) by simply reversing the sign of p_μ in (12.344):

$$\left(\gamma^\mu(-p_\mu) - mc\right)u_j(-p_\mu) = 0, \quad \Longrightarrow \quad \left(\gamma^\mu p_\mu + mc\right)v_j(p_\mu) = 0, \qquad j = 1,\ 2. \tag{12.346}$$

Example 12.7

(a) Verify that the *free particle* states $\psi_1(x^\mu)$ and $\psi_2(x^\mu)$ listed in (12.337) correspond to *positive energy*, $E^{(+)}$, and *positive momentum*, $\vec{p}$, solutions.

(b) Verify that, if we use the *normal* definitions of the energy and momentum operators (i.e., $\hat{H} = i\hbar\partial/\partial t$ and $\hat{\vec{p}} = -i\hbar\vec{\nabla}$), the *antiparticle* states $\Phi_3(x^\mu)$ and $\Phi_4(x^\mu)$ listed in (12.338) would correspond to *negative energy*, $-E^{(+)}$, and *negative momentum*, $-\vec{p}$, solutions.

(c) Verify that, if we use the *antiparticle* definitions of the energy and momentum operators (i.e., $\hat{H}^{(A)} = -i\hbar\partial/\partial t$ and $\hat{\vec{p}}^{\,(A)} = i\hbar\overrightarrow{\nabla}$), the *antiparticle* states $\Phi_3(x^\mu)$ and $\Phi_4(x^\mu)$ listed in (12.338) would lead to *positive energy*, $E^{(+)}$, and *positive momentum*, $\vec{p}$, solutions.

Solution

(a) First, we can easily show that $\psi_1(x)$ and $\psi_2(x)$ are eigenstates of the Dirac Hamiltonian with positive energy eigenvalues $E^{(+)}$. To see this, we only need to apply $\hat{H} = i\hbar\partial/\partial t$ on (12.337):

$$\hat{H}\psi_1(x^\mu) = i\hbar\frac{\partial}{\partial t}\left(u_1 e^{-i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) = E^{(+)}\left(u_1 e^{-i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) \equiv E^{(+)}\psi_1(x^\mu); \tag{12.347}$$

the same applies for $\psi_2(x)$.

Second, by applying the 3-momentum operator $\hat{\vec{p}} = -i\hbar\overrightarrow{\nabla}$ on (12.337), we can verify that $|\psi_1\rangle(x)$ and $|\psi_2\rangle(x)$ are eigenstates of the 3-momentum operator $\hat{\vec{p}}$ with a positive momentum $\vec{p}$; for instance, applying $\hat{p}_x = -i\hbar\partial/\partial x$ on (12.337), we have[45]

$$\hat{p}_x\,\psi_1(x^\mu) = -i\hbar\frac{\partial}{\partial x}\left(u_1 e^{-i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) = p_x\left(u_1 e^{-i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) \equiv p_x\,\psi_1(x^\mu), \tag{12.348}$$

which can be generalized to

$$\hat{\vec{p}}\,\psi_1(x^\mu) = -i\hbar\overrightarrow{\nabla}\left(u_1 e^{-i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) \equiv \vec{p}\,\psi_1(x^\mu); \tag{12.349}$$

the same applies for $\psi_2(x^\mu)$.

From (12.347) and(12.349), we conclude that the *free particle* states (12.337) do correspond to *positive energy*, $E^{(+)}$, and *positive 3-momentum*, $\vec{p}$, solutions.

(b) Applying the normal energy operator $\hat{H} = i\hbar\partial/\partial t$ on (12.338), we obtain

$$\hat{H}\Phi_3(x^\mu) = i\hbar\frac{\partial}{\partial t}\left(v_1 e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) = -E^{(+)}\left(v_1 e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) \equiv -E^{(+)}\Phi_3(x^\mu); \tag{12.350}$$

the same applies for $\Phi_4(x^\mu)$.

Similarly, applying the normal 3-momentum operator $\hat{p}_x = -i\hbar\partial/\partial x$ on (12.338), we have

$$\hat{p}_x\,\Phi_3(x^\mu) = -i\hbar\frac{\partial}{\partial x}\left(v_1 e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) = -p_x\left(v_1 e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) \equiv -p_x\,\Phi_3(x^\mu), \tag{12.351}$$

which can be generalized to

$$\hat{\vec{p}}\,\Phi_3(x^\mu) = -i\hbar\overrightarrow{\nabla}\left(v_1 e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) \equiv -\vec{p}\,\Phi_3(x^\mu); \tag{12.352}$$

the same applies for $\Phi_4(x^\mu)$.

From (12.350) and(12.352), it is easy to see that, when we use the *normal definitions* of the energy and momentum operators $\hat{H} = i\hbar\partial/\partial t$ and $\hat{\vec{p}} = -i\hbar\overrightarrow{\nabla}$, the *antiparticle* states (12.338) would correspond to solutions with *negative energy*, $-E^{(+)}$, and *negative momentum*, $-\vec{p}$.

[45] We should not confuse between the x in $\hat{p}_x$ and in $|\psi_1(x^\mu)\rangle$: the x in $\hat{p}_x$ refers to the x-component of the 3-momentum operator.

(c) Applying the energy operator corresponding to *antiparticles* $\hat{H}^{(A)} = -i\hbar\partial/\partial t$ on (12.338), we obtain

$$\hat{H}^{(A)}\Phi_3(x^\mu) = -i\hbar\frac{\partial}{\partial t}\left(v_1 e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) = E^{(+)}\left(v_1 e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) \equiv E^{(+)}\Phi_3(x^\mu); \tag{12.353}$$

the same applies for $\Phi_4(x^\mu)$.

Similarly, applying the 3-momentum operator corresponding to *antiparticles* $\hat{\vec{p}}^{\,(A)} = i\hbar\overrightarrow{\nabla}$ on (12.338), we can easily verify that $\Phi_3(x^\mu)$ and $\Phi_4(x^\mu)$ are eigenstates of the 3-momentum operator $\hat{\vec{p}}$ with a *positive momentum*:

$$\hat{\vec{p}}^{\,(A)}\,\Phi_3(x^\mu) = i\hbar\overrightarrow{\nabla}\left(v_1 e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}\right) \equiv \vec{p}\,\Phi_3(x^\mu); \tag{12.354}$$

the same applies for $\Phi_4(x^\mu)$.

From (12.353) and(12.354), we see that, when we use the *antiparticle* definitions of the energy and momentum operators $\hat{H}^{(A)} = -i\hbar\partial/\partial t$ and $\hat{\vec{p}}^{\,(A)} = i\hbar\overrightarrow{\nabla}$, the *antiparticle* states $\Phi_3(x^\mu)$ and $\Phi_4(x^\mu)$ listed in (12.338) do correspond to *positive energy*, $E^{(+)}$, and *positive momentum*, $\vec{p}$, solutions as postulated by the Feynman-Stüeckelberg interpretation.

Example 12.8

Show that the free *antiparticle* solutions (12.338) satisfy $\left(\gamma^\mu p_\mu + mc\right)v_j = 0$ with $j = 1, 2$.

Solution

To solve this problem, we need to calculate $\gamma^\mu\hat{p}_\mu\Phi_3(x^\mu)$. First, using (12.198), we can write

$$\begin{aligned}\gamma^\mu\hat{p}_\mu &= i\hbar\gamma^\mu\partial_\mu = \frac{i\hbar\gamma^0}{c}\frac{\partial}{\partial t} + i\hbar\vec{\gamma}\cdot\overrightarrow{\nabla} \\ &= \frac{i\hbar\gamma^0}{c}\frac{\partial}{\partial t} + i\hbar\gamma^1\frac{\partial}{\partial x} + i\hbar\gamma^2\frac{\partial}{\partial y} + i\hbar\gamma^3\frac{\partial}{\partial z}.\end{aligned} \tag{12.355}$$

Let us illustrate the calculations on $\Phi_3(x^\mu)$ only. As shown in (12.350) and (12.352), we have

$$\hat{H}\Phi_3(x^\mu) = -E^{(+)}\Phi_3(x^\mu), \tag{12.356}$$

and

$$\vec{\gamma}\cdot\hat{\vec{p}}\,\Phi_3(x^\mu) = -i\hbar\vec{\gamma}\cdot\overrightarrow{\nabla}\Phi_3(x^\mu) = -\left(\gamma^1 p_x + \gamma^2 p_y + \gamma^3 p_z\right)\Phi_3(x^\mu) \equiv -\vec{\gamma}\cdot\vec{p}\,\Phi_3(x^\mu). \tag{12.357}$$

Combining (12.355)–(12.357), we obtain

$$\begin{aligned}\gamma^\mu\hat{p}_\mu\Phi_3(x^\mu) &= \left(-\frac{\gamma^0 E^{(+)}}{c} - \gamma^1 p_x - \gamma^2 p_y - \gamma^3 p_z\right)\Phi_3(x^\mu) = -\left(\frac{\gamma^0 E^{(+)}}{c} + \vec{\gamma}\cdot\vec{p}\right)\Phi_3(x^\mu) \\ &\equiv -\gamma^\mu p_\mu\Phi_3(x^\mu).\end{aligned} \tag{12.358}$$

Inserting this result into Dirac's equation $\left(\gamma^\mu\hat{p}_\mu - mc\right)\Phi_3(x^\mu) = 0$ and simplifying the exponent, we end up with the following equation for the *antiparticle spinor* v_1

$$\left(\gamma^\mu p_\mu + mc\right)v_1 = 0. \tag{12.359}$$

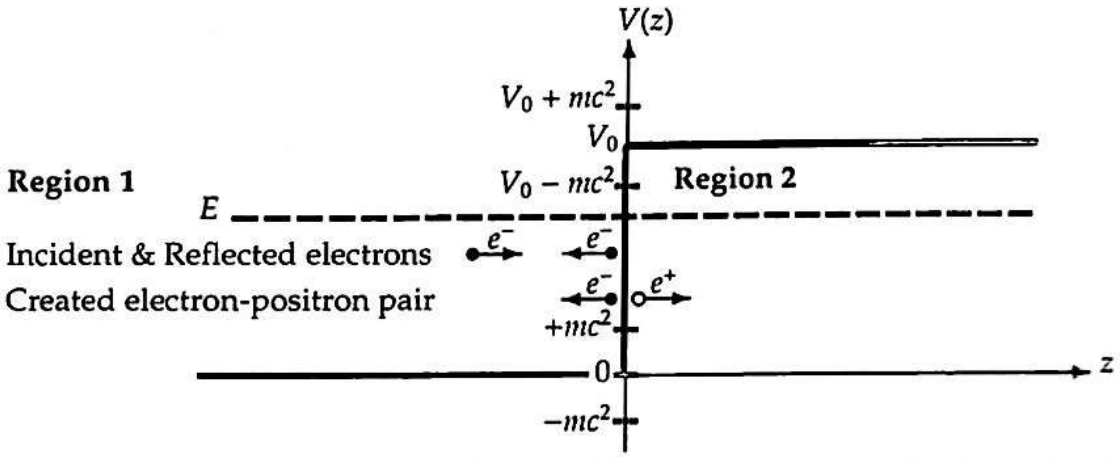


Figure 12.5: Scattering of electrons of energy E from a step potential V_0 where $E < V_0 - mc^2$.

Up to now, we have seen two ways of interpreting the negative energy solutions of the Dirac equation – the hole theory and the Feynman-Stückelberg interpretation. In the next section, we are going to look at a third method to interpret the negative energies; it will be provided by studying the scattering of a relativistic spin-1/2 particle from a electrostatic step potential.

12.4.12 Negative-Energies Interpretation: Klein Paradox for Spin-1/2 Particles

Another interpretation of the negative energy solutions is offered by the process of scattering a relativistic, spin-1/2 particle of energy E from a step potential of height V_0: $V(z) = 0$ for $z < 0$ and $V(z) = V_0$ for $z \geq 0$ as illustrated in Figure 12.5. We have solved this problem in Subsection 12.3.5 for a spinless particle using the Klein–Gordon equation; however, in this subsection, we are going to solve this problem using the Dirac equation. After carrying out the calculations below, we will see that the Dirac equation yields the same results as those of the Klein–Gordon equation obtained above except for spin statistics. The Dirac equation deals with fermions (spin-half particles) while the Klein–Gordon equation with bosons (spin-integer particles); so the Pauli exclusion principle needs to be taken into account in dealing with particle-antiparticle pair production in the scattering. As we will see below, the Dirac equation also encounters the Klein paradox and we will show how to resolve it using Dirac's hole theory by positing the creation of electron-positron pairs at the barrier.

Following the same approach outlined in Subsection 12.3.5, we are going to consider the following three energy cases.

Case $E > V_0 + mc^2$: Oscillating reflected and transmitted waves:
As seen in Problem (12.29) on Page 862, when $E > V_0 + mc^2$, we have both an oscillating reflected wave and an oscillating transmitted wave (see Eq. (12.850)), much like the case of a non-relativistic particle that we saw in Chapter 4 and the relativistic spinless particle we studied in Subsection 12.3.5.

Case $V_0 - mc^2 < E < V_0 + mc^2$: Total reflection:
When $V_0 - mc^2 < E < V_0 + mc^2$ (i.e., $-mc^2 < E_2 < mc^2$), the momentum in Region 2 of Figure 12.5 becomes *pure imaginary* (see Eq. (12.848)):

$$c^2p_2^2 = E_2^2 - m^2c^4 = (E - V_0)^2 - m^2c^4 < 0 \qquad \Longrightarrow \qquad p_2 = ip_2'; \tag{12.360}$$

where p_2' is real and positive. Similarly, the dimensionless factor λ listed in (12.855) also becomes imaginary

$$\lambda = \frac{p_2}{p_1}\frac{(E + mc^2)}{(E_2 + mc^2)} \qquad \Longrightarrow \qquad \lambda = i\lambda' = i\frac{p_2'}{p_1}\frac{(E + mc^2)}{(E - V_0 + mc^2)}, \tag{12.361}$$

where λ' is real and positive as well. Hence, *the transmitted wave function* (12.850) becomes an *exponentially decaying function*:

$$\psi_2(z) = N_T \begin{pmatrix} 1 \\ 0 \\ \frac{cp_2}{E_2+mc^2} \\ 0 \end{pmatrix} e^{ip_2z/\hbar} = N_T \begin{pmatrix} 1 \\ 0 \\ \frac{icp_2'}{E-V_0+mc^2} \\ 0 \end{pmatrix} e^{-p_2'z/\hbar}. \tag{12.362}$$

In this case, the transmitted current density becomes zero

$$\begin{aligned} j_T &= c\psi_T^\dagger \alpha_z \psi_T = c|N_T|^2 \begin{pmatrix} 1 & 0 & \frac{cp_2^*}{E-V_0+mc^2} & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{cp_2}{E-V_0+mc^2} \\ 0 \end{pmatrix} \\ &= \frac{2c^2(p_2 + p_2^*)}{E - V_0 + mc^2}|N_T|^2 \\ &= 0, \end{aligned} \tag{12.363}$$

and so does the transmission coefficient

$$T = \left|\frac{j_T}{j_I}\right| = 0. \tag{12.364}$$

The particle does not have sufficient energy to surmount the step potential. Hence, we have total reflection

$$R = \left|\frac{j_R}{j_I}\right| = 1, \tag{12.365}$$

because, as (12.869) indicates, the reflected and incident currents have equal magnitudes:

$$\left|\frac{j_R}{j_I}\right| = \left|-\frac{1-\lambda}{1+\lambda}\right|^2 = \left|\frac{1-i\lambda'}{1+i\lambda'}\right|^2 = \frac{1+\lambda'^2}{1+i\lambda'^2} = 1 \qquad \Longrightarrow \qquad |j_R| = |j_I|. \tag{12.366}$$

Additionally, as the particle tunnels through the potential's barrier, its probability density

decays exponentially:

$$\begin{aligned}\rho_T &= \psi_T^\dagger \psi_T = |N_T|^2 e^{-p_2' z/\hbar} \begin{pmatrix} 1 & 0 & \frac{cp_2'}{E-V_0+mc^2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{cp_2}{E-V_0+mc^2} \\ 0 \end{pmatrix} e^{-p_2' z/\hbar} \\ &= \left[1 + \frac{c^2|p_2|^2}{(E-V_0+mc^2)^2}\right] |N_T|^2 e^{-2p_2' z/\hbar} \\ &= \frac{2(E-V_0)}{E-V_0+mc^2} |N_T|^2 e^{-2p_2' z/\hbar}. \end{aligned} \tag{12.367}$$

This shows that the transmitted probability density can be positive or negative:

$$\text{If} \quad V_0 < E < V_0 + mc^2 \quad \Longrightarrow \quad \rho_T > 0, \tag{12.368}$$

$$\text{If} \quad V_0 - mc^2 < E < V_0 \quad \Longrightarrow \quad \rho_T < 0. \tag{12.369}$$

As we will see below, when the potential is sufficiently strong (i.e., when $E < V_0$), a particle-antiparticle pair is created at the barrier; particle-antiparticle pair creation occurs typically near a strong potential that changes abruptly over a very short distance.

Case $V_0 > E + mc^2$: Klein Paradox

If the barrier is too strong, $V_0 > E + mc^2$, the dimensionless parameter λ is negative (see Eq. (12.855)):

$$\lambda = \frac{p_2}{p_1} \frac{(E+mc^2)}{(E-V_0+mc^2)} < 0. \tag{12.370}$$

Hence, the *reflected coefficient is larger than 1* (see (12.869))

$$R = \left|\frac{j_R}{j_I}\right| = \left|-\left(\frac{1-\lambda}{1+\lambda}\right)^2\right| > 1, \tag{12.371}$$

and the *reflected current becomes larger than the incident current*:

$$\left|\frac{j_R}{j_I}\right| = \left|-\left(\frac{1-\lambda}{1+\lambda}\right)^2\right| > 1 \quad \Longrightarrow \quad |j_R| > |j_I|. \tag{12.372}$$

Note that, in addition to $|j_R| > |j_I|$ and $R > 1$, there is transmission into the classically forbidden region (i.e., Region 2); the transmission coefficient is non-zero (see (12.874)):

$$T = \left|\frac{J_T}{J_I}\right| = \left|\frac{4\lambda}{(1+\lambda)^2}\right| \neq 0, \tag{12.373}$$

and the probability density in Region 2 is positive and also non-zero:

$$\begin{aligned}\rho_T &= \psi_T^\dagger \psi_T = |N_T|^2 e^{ip_2 z/\hbar} \begin{pmatrix} 1 & 0 & \frac{cp_2'}{E-V_0+mc^2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{cp_2}{E-V_0+mc^2} \\ 0 \end{pmatrix} e^{-ip_2 z/\hbar} \\ &= \frac{2(E-V_0)}{E-V_0+mc^2} |N_T|^2 \neq 0. \end{aligned} \tag{12.374}$$

Clearly, much like what we have encountered in Subsection 12.3.5 where we treated the scattering of spin-zero particles from a step potential, the Klein paradox (i.e., $R > 1$) also occurs in the scattering of spin-1/2 particles from a strong potential. In what follows, we are going to provide an interpretation of the Klein paradox.

Interpretation of the Klein paradox using Dirac's hole theory
The only way to resolve the klein paradox is by going beyond the single particle nature of the Dirac equation and considering the creation of particle-antiparticle pairs at the barrier; due to the "violent" interaction of the incident flux with a strong potential that changes abruptly over a very short distance, particle–antiparticle pairs get created from the vacuum (Dirac sea). In order for a particle-antiparticle pair to be created at the barrier, the barrier must have a minimum size of $V_{0_{min}} = 2mc^2$ or $V_0 > E + mc^2 > 2mc^2$. To see this, notice that if the kinetic energies of the particle and the antiparticle are both zero, the energy of a particle on the left side of the barrier ($z < 0$) is mc^2 and that of an antiparticle on the right side of the barrier ($z > 0$) is $mc^2 - V_0$; hence, an energy gap of $2mc^2$ between them. So, due to the strong potential, the incident electrons striking the potential barrier will stimulate the creation of electron-positron pairs at the barrier; that is, each incident electron striking the potential barrier will stimulate the excitation of a positive energy electron from the negative energy sea (i.e., $E < -mc^2$), leaving behind a positive energy *hole* — a positron — such that the positive energy electron travels to the left ($z < 0$) and the positive energy positron travels to the right ($z > 0$). As illustrated in Figure 12.5, the electrons knocked out of the negative energy sea end up adding to the reflected electron beam in Region 1 to create a total reflected current that is larger than the incident current. Now, how to interpret negative probabilities such as the result displayed in Eq. (12.369)? They can be viewed as follows: the negative probability density for finding a particle in Region 2 is equivalent to a positive probability density for finding an antiparticle.

Klein paradox and the negative energy solutions of the Dirac equation
What is the connection between the Klein paradox and Dirac's negative energy solutions? First, as a reminder, one cannot exclude the negative-energy states listed in Eq. (12.338) since the two positive-energy states (12.337) cannot represent by themselves a complete set of solutions; we need four solutions: two corresponding to spin-up and spin-down positive energy states, and two corresponding to spin-up and spin-down negative energy states. Second, and as seen above, the Klein paradox provides an adequate explanation for the existence/creation of electron-positron pairs during the process of scattering electron beams from strongly repulsive electric potentials. Since the positive energy solutions of the Dirac equation describe electron states (i.e., particles), the negative energy solutions can only correspond to positrons (i.e., antiparticles); *the negative energy solutions correspond to positive energy antiparticles that travel forwards in time.* As illustrated in Figure 12.5, the positron that emerges from the created electron-positron pair travels to the right in Region 2.

In summary, the two studies we carried out in Subsections 12.3.5 and 12.4.12 demonstrate the insufficiency of the Klein–Gordon and Dirac equations. In describing the scattering of spin-zero and spin-half single particles from strong step potentials, both equations lead to the Klein paradox. Both equations started as single particle prescriptions and ended up dealing forcibly with multiparticle systems — particle-antiparticle pair creation — so as to make sense of the anomalies encountered; the klein paradox cannot be completely resolved until we consider the creation of particle-antiparticle pairs. By forcing or imposing the single particle requirement on these two equations, we inevitably end up reaping various anomalies

and inconsistencies.

Clearly, the only way out is to go beyond the single-particle limitation and consider the solutions of the Klein–Gordon and Dirac equations, Ψ, as describing *fields*, not single-particle wave functions. In this case, the squared moduli of the fields, $\Psi^*\Psi$, need to be multiplied by charge to make them *charge densities*, not quantum mechanical wave functions whose squared moduli are probability densities; that is, we need to work with *fields*, Ψ, such that $\rho = e\Psi^*\Psi$ represents a *charge density*, and not with single-particle wave functions Ψ where $\rho = \Psi^*\Psi$ represents a probability density. As we will see at the end of this chapter (see Section 12.5), relativistic phenomena cannot be treated by means of single-particle quantum mechanics, but rather within the context of *field theories* as they are equipped with the theoretical tools to handle multiparticle systems such as particle-antiparticle pair creation and annihilation.

Summary of the Interpretation of the Dirac Negative Energy Solutions

In the hole theory, Dirac viewed a negative energy state as describing a positive energy electron that is created from the vacuum and which is accompanied by the appearance of a hole with a negative energy. Dirac postulated that the hole is the antiparticle of the electron and called it the positron, for it has the same mass as the electron, but with an opposite charge $+e$.

In the Feynman-Stückelberg interpretation, the negative energy solution is viewed as describing either a negative energy particle traveling backwards in time or, equivalently, a positive energy antiparticle traveling forwards in time (because a negative energy particle traveling backwards in time is mathematically equivalent to a positive energy antiparticle traveling forwards in time as shown in Eq. (12.328)).

The Klein paradox studied in Subsections 12.3.5 and 12.4.12 has provided us with another interpretation of the negative energy solutions as originating from the creation of particle-antiparticle pairs. It is important to note that the appearance of negative energy solutions in the Klein Gordon and Dirac equations have served as *the main catalyst that led to the discovery of antiparticles*.

12.4.13 Spin-Half Solutions of the Dirac Equation

In addition to the conclusions reached above that the Dirac spinors describe spin-half particles (see Page 759 and Eqs (12.337)–(12.338)), we are going to further our study of the spin in this section. We are going to consider another, yet more explicit, proof that the Dirac equation describes spin 1/2 particles. For this, we need to find an operator that commutes with the Dirac Hamiltonian (i.e., a constant of the motion) and that its eigenvalues are $\pm\hbar/2$.

First, let us find out if the orbital angular momentum $\hat{\vec{L}}$ commutes with the Dirac Hamiltonian $\hat{H} = c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2$. Using equation (12.148), we can write

$$[\hat{H},\ \hat{\vec{L}}] = \left[\left(c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2\right),\ \hat{\vec{L}}\right] = \left[\left(c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2\right),\ \hat{\vec{r}}\times\hat{\vec{p}}\right] \equiv c\left[\vec{\alpha}\cdot\hat{\vec{p}},\ \hat{\vec{r}}\times\hat{\vec{p}}\right], \tag{12.375}$$

where we have used the fact the constant factor βmc^2 commutes with $\hat{\vec{L}}$ in the last step of the above equation. To find an expression for this commutator, let us begin by calculating its x component; using the relation $\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y$, we have

$$[\hat{H},\ \hat{L}_x] = c\left[\vec{\alpha}\cdot\hat{\vec{p}},\ \hat{y}\hat{p}_z - \hat{z}\hat{p}_y\right] = c\left[\alpha_x\hat{p}_x + \alpha_y\hat{p}_y + \alpha_z\hat{p}_z,\ \hat{y}\hat{p}_z - \hat{z}\hat{p}_y\right]. \tag{12.376}$$

The only terms that survive are those involving $[\hat{y},\ \hat{p}_y] = [\hat{z},\ \hat{p}_z] = i\hbar$:

$$\begin{aligned}[\hat{H},\ \hat{L}_x] &= c\left[\alpha_x\hat{p}_x + \alpha_y\hat{p}_y + \alpha_z\hat{p}_z,\ \hat{y}\hat{p}_z - \hat{z}\hat{p}_y\right] = c\left[\alpha_y\hat{p}_y,\ \hat{y}\hat{p}_z\right] - c\left[\alpha_z\hat{p}_z,\ \hat{z}\hat{p}_y\right] \\ &= c\alpha_y\left[\hat{p}_y,\ \hat{y}\right]\hat{p}_z - c\alpha_z\left[\hat{p}_z,\ \hat{z}\right]\hat{p}_y \\ &= -ic\hbar\alpha_y\hat{p}_z + ic\hbar\alpha_z\hat{p}_y \\ &\equiv -ic\hbar\left(\alpha_y\hat{p}_z - \alpha_z\hat{p}_y\right). \end{aligned} \tag{12.377}$$

Since the term $(-\alpha_y\hat{p}_z + \alpha_z\hat{p}_y)$ is nothing but the $x-$component of the cross product between $\vec{\alpha}$ and $\hat{\vec{p}}$, we can write

$$[\hat{H},\ \hat{L}_x] = -ic\hbar\alpha_y\hat{p}_z + ic\hbar\alpha_z\hat{p}_y = -ic\hbar\left(\vec{\alpha}\times\hat{\vec{p}}\right)_x. \tag{12.378}$$

Now, using the same method (or simply taking the circular permutations of xyz components in (12.378)), we can infer the following two relations

$$[\hat{H},\ \hat{L}_y] = -ic\hbar\alpha_z\hat{p}_x + ic\hbar\alpha_x\hat{p}_z \equiv -ic\hbar\left(\vec{\alpha}\times\hat{\vec{p}}\right)_y, \tag{12.379}$$

$$[\hat{H},\ \hat{L}_z] = -ic\hbar\alpha_x\hat{p}_y + ic\hbar\alpha_y\hat{p}_x \equiv -ic\hbar\left(\vec{\alpha}\times\hat{\vec{p}}\right)_z. \tag{12.380}$$

Adding equations (12.378), (12.379), and (12.380), we obtain

$$\boxed{[\hat{H},\ \hat{\vec{L}}] = -i\hbar c(\ \vec{\alpha}\times\hat{\vec{p}}\ ),} \tag{12.381}$$

where $\hat{\vec{p}} = -i\hbar\vec{\nabla}$. This clearly shows that the Dirac Hamiltonian does not commute with the orbital angular momentum $\hat{\vec{L}}$. Hence, *the orbital angular momentum $\hat{\vec{L}}$ is not a constant of the motion.*

Second, let us find out if the Dirac Hamiltonian commutes with the following 4×4 operator

$$\hat{\vec{S}} = \frac{\hbar}{2}\hat{\vec{\Sigma}} = \frac{\hbar}{2}\begin{pmatrix}\vec{\sigma} & 0\\ 0 & \vec{\sigma}\end{pmatrix} \tag{12.382}$$

where $\vec{\sigma}$ are the 2×2 Pauli matrices. We can write the $xyz-$components of $\hat{\vec{S}}$ as

$$S_x = \frac{\hbar}{2}\begin{pmatrix}\sigma_x & 0\\ 0 & \sigma_x\end{pmatrix},\qquad S_y = \frac{\hbar}{2}\begin{pmatrix}\sigma_y & 0\\ 0 & \sigma_y\end{pmatrix},\qquad S_z = \frac{\hbar}{2}\begin{pmatrix}\sigma_z & 0\\ 0 & \sigma_z\end{pmatrix}. \tag{12.383}$$

Inserting the 2×2 matrices (12.162) into (12.383), we obtain

$$S_x = \frac{\hbar}{2}\begin{pmatrix}0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{pmatrix},\quad S_y = \frac{\hbar}{2}\begin{pmatrix}0 & -i & 0 & 0\\ i & 0 & 0 & 0\\ 0 & 0 & 0 & -i\\ 0 & 0 & i & 0\end{pmatrix},\quad S_z = \frac{\hbar}{2}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}. \tag{12.384}$$

In Problem 12.23 (see Equation (12.815)), we have shown that the Dirac Hamiltonian does not commute with $\hat{\vec{\Sigma}}$: $[\hat{H},\ \hat{\vec{\Sigma}}] = 2ic(\vec{\alpha}\times\hat{\vec{p}})$. Hence, we can write

$$\boxed{[\hat{H},\ \hat{\vec{S}}] = i\hbar c(\ \vec{\alpha}\times\hat{\vec{p}}\ ).} \tag{12.385}$$

This shows that the *operator* $\hat{\vec{S}}$ *is not a constant of the motion.*

It is important to mention that, as shown above, although $\hat{\vec{L}}$ and $\hat{\vec{S}}$, taken separately, are not constants of the motion, their sum

$$\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}} \equiv \hat{\vec{L}} + \frac{\hbar}{2}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}, \tag{12.386}$$

is a constant of the motion; this can be easily shown by simply adding equations (12.381) and (12.385):

$$\boxed{[\hat{H}, \hat{\vec{J}}] = [\hat{H}, \hat{\vec{L}} + \hat{\vec{S}}] = 0.} \tag{12.387}$$

As shown in Problem 12.24 on Page 856 (see (12.819)), the operators $\hat{S}_x$, $\hat{S}_y$, $\hat{S}_z$ obey the standard commutation relations of a spin operator

$$\boxed{[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z, \qquad [\hat{S}_y, \hat{S}_z] = i\hbar\hat{S}_x, \qquad [\hat{S}_z, \hat{S}_x] = i\hbar\hat{S}_y.} \tag{12.388}$$

Additionally, $\hat{S}_z$ and $\vec{S}^2$ are diagonal (see equations (12.384) and (12.819)):

$$\vec{S}^2 = \frac{3}{4}\hbar^2\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{12.389}$$

which can be written as

$$\boxed{\vec{S}^2 = \tfrac{1}{2}\left(\tfrac{1}{2}+1\right)\hbar^2 I_4 \equiv S(S+1)\hbar^2 I_4,} \tag{12.390}$$

where $S = 1/2$ and I_4 is the 4×4 unity matrix. Hence, since $\vec{S}^2$ is proportional to the unity matrix, it is easy to verify that the Dirac spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$ listed in equations (12.321), (12.334), (12.335) are eigenstates of $\vec{S}^2$ with eigenvalue $3\hbar^2/4$:

$$\vec{S}^2|u_j\rangle = \frac{3}{4}\hbar^2|u_j\rangle, \qquad \vec{S}^2|v_j\rangle = \frac{3}{4}\hbar^2|v_j\rangle, \qquad j = 1, 2. \tag{12.391}$$

We should note that, in general, as shown in Problem 12.25 on Page 780, the Dirac spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$ are not eigenstates of $\hat{S}_z$. However, and as shown in Example 12.9 on Page 780, *for a particle that is traveling in the z–direction with a momentum* $\vec{p} = (0, 0, \pm p)$, *the Dirac spinors are eigenstates of* $\hat{S}_z$ *with eigenvalues* $\pm\hbar/2$:

$$\hat{S}_z|u_1\rangle = \frac{\hbar}{2}|u_1\rangle, \qquad \hat{S}_z|u_2\rangle = -\frac{\hbar}{2}|u_2\rangle, \tag{12.392}$$

$$\hat{S}_z^{(A)}|v_1\rangle = -\hat{S}_z|v_1\rangle = \frac{\hbar}{2}|v_1\rangle, \qquad \hat{S}_z^{(A)}|v_2\rangle = -\hat{S}_z|v_2\rangle = -\frac{\hbar}{2}|v_2\rangle. \tag{12.393}$$

Note the change of the sign of $\hat{S}_z$ when it acts on the antiparticle spinors $|v_1\rangle$ and $|v_2\rangle$. By analogy to (12.343) where we have seen that the orbital angular momentum for antiparticles

is equal to $\hat{\vec{L}}^{(A)} = -\hat{\vec{L}}$ and in order to insure that the total angular momentum is conserved (i.e., $[\hat{H}, \hat{\vec{L}} + \hat{\vec{S}}] = 0$), the spin for antiparticles $\hat{\vec{S}}^{(A)}$ must be equal to $-\hat{\vec{S}}$ for particles (i.e., the spin flips for antiparticles); in particular, for the $z-$component we have

$$\hat{S}_z^{(A)} = -\hat{S}_z. \tag{12.394}$$

We can summarize the above derived expressions involving the operator $\hat{\vec{S}}$ as follows

- Equation (12.387) shows that $\hat{\vec{S}}$ must be added to the orbital angular momentum $\hat{\vec{L}}$ so as to obtain a total angular momentum $\hat{\vec{J}}$ that is a constant of the motion: $[\hat{H}, \hat{\vec{L}} + \hat{\vec{S}}] = 0$,
- Equation (12.388) shows that the $xyz-$components of $\hat{\vec{S}}$ obey the standard commutation relations of an angular momentum operator: $[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z$, $[\hat{S}_y, \hat{S}_z] = i\hbar\hat{S}_x$, $[\hat{S}_z, \hat{S}_x] = i\hbar\hat{S}_y$,
- Equation (12.391) shows that the particle and antiparticle solutions of the Dirac equation (i.e., the spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$) are eigenstates of $\hat{\vec{S}}^2$ with eigenvalue $3\hbar^2/4$ very much like a spin-half particle: $\vec{S}^2|\psi_D\rangle = (3\hbar^2/4)|\psi_D\rangle$ where $|\psi_D\rangle$ is the Dirac wave function (either the particle states $|\psi_1\rangle$ and $|\psi_2\rangle$ or antiparticle states $|\Phi_3\rangle$ and $|\Phi_4\rangle$ listed in Eqs. (12.337)–(12.338)),
- Equations (12.392) and (12.393) show that the Dirac spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$ for a particle moving along the $z-axis$ are eigenstates of $\hat{S}_z$ with eigenvalues $\pm\hbar/2$ very much like a spin-half particle: $\hat{S}_z|\psi_D\rangle = \pm(\hbar/2)|\psi_D\rangle$.

These four observations lead to an unmistakable conclusion: *the solutions of the Dirac equation for a free particle traveling along the $z-$direction describe spin-half particles.* The quantity $\hat{\vec{S}}$ has all the ingredients of the standard spin-half angular momentum operator in quantum mechanics[46]. Consequently, one concludes that *the Dirac equation describes spin-half particles.* Thus, the spin is intrinsically embedded in a natural way and from the very outset in the formalism. Unlike the Schrödinger equation where the spin had to be introduced in an adhoc manner (i.e., by hand), the spin degree of freedom is naturally built into the Dirac formalism.

Example 12.9
Using equations (12.321), (12.334), (12.335) and (12.383), show that the Dirac spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$ are eigenstates of the operator $\hat{S}_z$ when the particle is traveling in the $z-$axis with a momentum $p_z = \pm|p\hat{k}| = \pm p$ where $p > 0$.

Solution

[46] From quantum mechanics, we know that the Pauli matrices (when multiplied by $\hbar/2$) represent spin-1/2 operators; this is the case here, since equation (12.382) shows that the operator $\hat{\vec{S}}$ is directly proportional to the Pauli matrices $\vec{sigma}$.

For a particle that is traveling in the z–direction, inserting $p_z = \pm p$ and $p_x = p_y = 0$ in the Dirac spinors (12.321), (12.334) and (12.335), we obtain

$$u_1(p) = N\begin{pmatrix} 1 \\ 0 \\ \frac{\pm cp}{E^{(+)}+mc^2} \\ 0 \end{pmatrix}, \quad u_2(p) = N\begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{\mp cp}{E^{(+)}+mc^2} \end{pmatrix}, \quad v_1(p) = N\begin{pmatrix} 0 \\ \frac{\mp cp}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix}, \quad v_2(p) = N\begin{pmatrix} \frac{\pm cp}{E^{(+)}+mc^2} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \tag{12.395}$$

where $E^{(+)} = \sqrt{c^2p^2 + m^2c^4}$.

Using the expression (12.395) of $|u_1\rangle$ and the matrix of $\hat{S}_z$ listed in (12.383), we have

$$\hat{S}_z|u_1\rangle = N\frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ \frac{\pm cp}{E^{(+)}+mc^2} \\ 0 \end{pmatrix} = N\frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \\ \frac{\pm cp}{E^{(+)}+mc^2} \\ 0 \end{pmatrix} = \frac{\hbar}{2}|u_1\rangle. \tag{12.396}$$

This equation clearly shows that $|u_1\rangle$ is an eigenstate of $\hat{S}_z$ with an eigenvalue of $\hbar/2$.

Following the same method and using equations (12.395) and (12.383), we can verify that $|u_2\rangle$ is also an eigenstate of $\hat{S}_z$ with an eigenvalue of $-\hbar/2$:

$$\hat{S}_z|u_2\rangle = N\frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{\mp cp}{E^{(+)}+mc^2} \end{pmatrix} = N\frac{\hbar}{2}\begin{pmatrix} 0 \\ -1 \\ 0 \\ \frac{\pm cp}{E^{(+)}+mc^2} \end{pmatrix} = -N\frac{\hbar}{2}\begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{\mp cp}{E^{(+)}+mc^2} \end{pmatrix} = -\frac{\hbar}{2}|u_2\rangle. \tag{12.397}$$

Now, let us look at the action of $\hat{S}_z$ on the *antiparticle* spinors $|v_1\rangle$ and $|v_2\rangle$; as explained above, the sign of $\hat{S}_z$ changes sign when we deal with antiparticle spinors. That is, using (12.394), we can write

$$\hat{S}_z^{(A)}|v_1\rangle = -\hat{S}_z|v_1\rangle = -N\frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ \frac{\mp cp}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix} = N\frac{\hbar}{2}\begin{pmatrix} 0 \\ \frac{\mp cp}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix} = \frac{\hbar}{2}|v_1\rangle. \tag{12.398}$$

clearly, $|v_1\rangle$ is an eigenstate of $\hat{S}_z^{(A)}$ with an eigenvalues of $\hbar/2$. Similarly, we can verify that $|v_2\rangle$ is an eigenstate of $\hat{S}_z^{(A)}$ with and eigenvalue of $-\hbar/2$:

$$\hat{S}_z^{(A)}|v_2\rangle = -\hat{S}_z|v_2\rangle = -N\frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} \frac{\pm cp}{E^{(+)}+mc^2} \\ 0 \\ 1 \\ 0 \end{pmatrix} = -N\frac{\hbar}{2}\begin{pmatrix} \frac{\pm cp}{E^{(+)}+mc^2} \\ 0 \\ 1 \\ 0 \end{pmatrix} = -\frac{\hbar}{2}|v_2\rangle. \tag{12.399}$$

Hence, we conclude that, for a particle which is traveling in the z–axis, the Dirac spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$ are eigenstates of the operator $\hat{S}_z$ with eigenvalues $\pm\hbar/2$, very much like a spin-half particle.

12.4.14 Helicity

The particle spinors u_1 and u_2 and the antiparticle spinors v_1 and v_2 are two-fold degenerate (see Eq. (12.336))

$$\hat{H}|u_j\rangle = E^{(+)}|u_j\rangle, \qquad \hat{H}|v_j\rangle = E^{(+)}|v_j\rangle, \quad j = 1,\ 2, \qquad E^{(+)} = \sqrt{c^2\vec{p}^{\,2} + m^2c^4}, \tag{12.400}$$

and, hence, the wave functions (12.337) and (12.338) are not completely specified; the spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$ require another degree of freedom that has yet to be specified. So, to be able to distinguish between these degenerate states, we would need an observable whose operator must commute with both of the Hamiltonian $\hat{H}$ and the momentum $\hat{\vec{p}}$, bearing in mind that we know already that $[\hat{H}, \hat{\vec{p}}\,] = 0$. This observable can neither be the orbital angular momentum nor the spin, since, as shown in equations (12.381) and (12.385), the Hamiltonian commutes neither with $\hat{\vec{L}}$ nor with $\hat{\vec{S}}$ separately. For a spin $S = 1/2$ particle, however, we can construct this observable in terms of the spin and the momentum; it is known as the *helicity* and it is defined as the component of the spin along the direction of the particle's motion (i.e., the projection of the spin along $\vec{p}$):

$$\widehat{\mathcal{H}} = \hat{\vec{S}}\cdot\hat{n} = \frac{\hbar}{2}\hat{\vec{\Sigma}}\cdot\hat{n} = \frac{\hbar}{2}\begin{pmatrix}\vec{\sigma} & 0\\ 0 & \vec{\sigma}\end{pmatrix}\cdot\hat{n} = \frac{\hbar}{2}\begin{pmatrix}\vec{\sigma}\cdot\hat{n} & 0\\ 0 & \vec{\sigma}\cdot\hat{n}\end{pmatrix}, \qquad \text{where} \qquad \hat{n} = \frac{\hat{\vec{p}}}{|\vec{p}|}\ ; \tag{12.401}$$

the quantity $\hat{n} = \hat{\vec{p}}/|\vec{p}|$ *is the unit vector in the direction of the particle's momentum* where $\hat{\vec{p}}$ is the momentum operator and $\vec{p}$ is just the momentum vector. We can also obtain the matrix representation of the helicity (this calculation was carried out in Problem 12.21 on Page 851):

$$\boxed{\widehat{\mathcal{H}} = \frac{\hbar}{2|\vec{p}|}\begin{pmatrix} \hat{p}_z & \hat{p}_x - i\hat{p}_y & 0 & 0\\ \hat{p}_x + i\hat{p}_y & -\hat{p}_z & 0 & 0\\ 0 & 0 & \hat{p}_z & \hat{p}_x - i\hat{p}_z\\ 0 & 0 & \hat{p}_x + i\hat{p}_y & -\hat{p}_z \end{pmatrix}.} \tag{12.402}$$

Eigenvalues of the Helicity Operator: left-handed and right-handed helicities
To find the eigenvalues of the helicity operator, it is useful to note that its square is given by $(1/4)I_4$, where I_4 is the 4×4 unit matrix. This can be shown easily as follows. As seen in Chapter 5, for any two vectors $\vec{A}$ and $\vec{B}$ which commute with the Pauli matrices $\vec{\sigma}$, we have

$$(\vec{\sigma}\cdot\vec{A})(\vec{\sigma}\cdot\vec{B}) = (\vec{A}\cdot\vec{B})I_2 + i\vec{\sigma}\cdot(\vec{A}\times\vec{B}), \tag{12.403}$$

where I_2 is the 2×2 unit matrix. Since the momentum operator $\hat{\vec{p}}$ commutes with the spin and since $\vec{p}\times\vec{p} = 0$, this relation leads to

$$(\vec{\sigma}\cdot\vec{p})(\vec{\sigma}\cdot\vec{p}) = (\vec{p}\cdot\vec{p})I_2 + i\vec{\sigma}\cdot(\vec{p}\times\vec{p}) \equiv \vec{p}^{\,2}I_2 \qquad \Longrightarrow \qquad (\vec{\Sigma}\cdot\hat{n}\,)(\vec{\Sigma}\cdot\hat{n}\,) = \hat{n}^{\,2}I_4 = I_4. \tag{12.404}$$

Using this relation, we can obtain the square of the helicity:

$$\widehat{\mathcal{H}}^{\,2} = \frac{\hbar^2}{4}\left(\hat{\vec{\Sigma}}\cdot\hat{n}\,\right)\left(\hat{\vec{\Sigma}}\cdot\hat{n}\,\right) = \frac{\hbar^2}{4}(\hat{n})^2\ I_4 = \frac{\hbar^2}{4}I_4. \tag{12.405}$$

(a) (b)

Figure 12.6: The two helicity states: **(a): Right-handed helicity:** with helicity eigenvalue $\hbar/2$, the spin is aligned in the same direction as the momentum. **(b): Left-handed helicity:** with helicity eigenvalue $-\hbar/2$, the spin is aligned in the opposite direction as the momentum.

We can also derive this result by using the matrix representation of the helicity; this calculation is carried out in Problem 12.22 on Page 853. Since $\widehat{\mathcal{H}}^2 = (\hbar^2/4)I_4$, the eigenvalues of the helicity operator for a spin-1/2 particle are $\pm\hbar/2$. What do these two eigenvalues represent? They correspond to the two possible helicity states of particles and antiparticles. As illustrated in Figure 12.6, the helicity eigenvalue $\hbar/2$, where the spin vector $\vec{S}$ is aligned in the same direction as the momentum vector $\vec{p}$, corresponds to what is called the *right-handed helicity* eigenstate and the $-\hbar/2$ eigenvalue, where the spin is aligned in the opposite direction as the momentum, corresponds to the *left-handed helicity* eigenstate. So, *the "handedness" of a particle has to do with the relative orientation of its spin and momentum.* As we will see below, for a massless particle, the helicity is Lorentz invariant, since the helicity scalar product is invariant under Lorentz transformations.

12.4.14.1 Joint Eigenstates of the Hamiltonian, Momentum, and Helicity Operators

To be able to obtain free particle wave functions that can be simultaneously eigenstates of the Hamiltonian, the momentum, and the helicity, we would need to show that $\widehat{\mathcal{H}}$ commutes with $\hat{H}$ and $\hat{\vec{p}}$. First, since the momentum operator commutes with itself and with the spin operator, $[\hat{\vec{S}}, \hat{\vec{p}}] = 0$, and since $\widehat{\mathcal{H}} = \hat{\vec{S}} \cdot \hat{n} = \hat{\vec{S}} \cdot \hat{\vec{p}}/|\vec{p}|$, it is easy to see that the helicity commutes with the momentum

$$[\widehat{\mathcal{H}}, \hat{\vec{p}}] \equiv \left[\hat{\vec{S}} \cdot \frac{\hat{\vec{p}}}{|\vec{p}|}, \hat{\vec{p}}\right] = \frac{1}{|\vec{p}|}[\hat{\vec{S}} \cdot \hat{\vec{p}}, \hat{\vec{p}}] = \frac{1}{|\vec{p}|}[\hat{S}_j \hat{p}_j, \hat{\vec{p}}] = \frac{1}{|\vec{p}|}\left(\hat{S}_j[\hat{p}_j, \hat{\vec{p}}] + [\hat{S}_j, \hat{\vec{p}}]\hat{p}_j\right) = 0 \quad (12.406)$$

where the Einstein summation over the index $j = 1, 2, 3$ is used. Second, and as shown in Problem 12.21 (see Eq. (12.799)), the Hamiltonian also commutes with the helicity operator

$$\boxed{[\hat{H}, \widehat{\mathcal{H}}] = \left[c\vec{\alpha} \cdot \hat{\vec{p}} + \beta mc^2, \hat{\vec{S}} \cdot \hat{n}\right] = 0.} \quad (12.407)$$

The eigenvalues of the helicity operator are then constants of the motion and, hence, good quantum numbers to label the corresponding wave functions.

Now, since $\hat{H}$, $\hat{\vec{p}}$ and $\widehat{\mathcal{H}}$ mutually commute, we can find a set of wave functions that are simultaneously eigenstates of the Hamiltonian, the momentum, and the helicity. For a free

particle that is moving in an arbitrary direction $(\theta,\ \phi)$, we have shown in Problem 12.35 (see Eqs. (12.1001)–(12.1002)) that the particle and anti-particle helicity as well as Hamiltonian eigenstates are given by

$$\boxed{\psi_+(x^\mu) = N\begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) e^{i\phi} \\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right) e^{i\phi} \end{pmatrix} e^{-i(Et-\vec{p}\cdot\vec{r})}, \quad \psi_-(x^\mu) = N\begin{pmatrix} \sin\left(\frac{\theta}{2}\right) \\ -\cos\left(\frac{\theta}{2}\right) e^{i\phi} \\ -\frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right) e^{i\phi} \end{pmatrix} e^{-i(Et-\vec{p}\cdot\vec{r})},} \tag{12.408}$$

$$\boxed{\Phi_+(x^\mu) = N\begin{pmatrix} -\frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right) e^{i\phi} \\ \sin\left(\frac{\theta}{2}\right) \\ -\cos\left(\frac{\theta}{2}\right) e^{i\phi} \end{pmatrix} e^{i(Et-\vec{p}\cdot\vec{r})}, \quad \Phi_-(x^\mu) = N\begin{pmatrix} \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right) e^{i\phi} \\ \cos(\theta/2) \\ \sin\left(\frac{\theta}{2}\right) e^{i\phi} \end{pmatrix} e^{i(Et-\vec{p}\cdot\vec{r})},} \tag{12.409}$$

where $E = \sqrt{\vec{p}^{\,2}c^2 + m^2c^4}$ and the normalization constant N is given by $N = \sqrt{E+mc^2}$ (see Eq. (12.952)). In particular, for a free particle moving along the positive $z-$axis direction (i.e., $\theta = 0$ or $\hat{n} = (0,\ 0,\ +1)$), these eigenstates become

$$\psi_+(x^\mu) = N\begin{pmatrix} 1 \\ 0 \\ \frac{cp}{E^{(+)}+mc^2} \\ 0 \end{pmatrix} e^{i(Et-pz)}, \qquad \psi_-(x^\mu) = N\begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{-cp}{E^{(+)}+mc^2} \end{pmatrix} e^{i(Et-pz)}, \tag{12.410}$$

$$\Phi_+(x^\mu) = N\begin{pmatrix} 0 \\ \frac{-cp}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix} e^{i(Et-pz)}, \qquad \Phi_-(x^\mu) = N\begin{pmatrix} \frac{cp}{E^{(+)}+mc^2} \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{i(Et-pz)}, \tag{12.411}$$

12.4.15 Dirac Equation for Massless Particles: Weyl Fermions

An interesting application of the Dirac equation is to study its solutions for *massless* particles. In particular, we want to focus on the violation of parity by the solutions of the massless Dirac equation and the connection of these solutions to neutrinos.

To obtain the solutions of the Dirac equation for a free, *massless* particle, we would need to insert $m = 0$ in the coupled Dirac equations (12.708) and (12.709); this leads to

$$E\,u_A = c\vec{\sigma}\cdot\vec{p}\,u_B, \tag{12.412}$$

$$E\,u_B = c\vec{\sigma}\cdot\vec{p}\,u_A. \tag{12.413}$$

These are known as the (coupled) *Weyl equations*[47]. Adding and subtracting these two equations, we obtain

$$E\,(u_A + u_B) = c\vec{\sigma}\cdot\vec{p}\,(u_A + u_B), \tag{12.414}$$

$$E\,(u_A - u_B) = -c\vec{\sigma}\cdot\vec{p}\,(u_A - u_B). \tag{12.415}$$

[47] Hermann Weyl, *Electron and gravitation*, Z. Phys. **56**, 330—352 (1929). An English translation can be found in: Surveys High Energ. Phys. **5**, 261—267 (1986).

By introducing the following states

$$\phi_R = \frac{1}{2}(u_A + u_B), \qquad \phi_L = \frac{1}{2}(u_A - u_B), \tag{12.416}$$

it is easy to rewrite equations (12.414) and (12.415) in the following form of a set of *decoupled* equations:

$$E\,\phi_R = c\vec{\sigma}\cdot\vec{p}\,\phi_R, \tag{12.417}$$
$$E\,\phi_L = -c\vec{\sigma}\cdot\vec{p}\,\phi_L; \tag{12.418}$$

the subscripts R and L will be clarified shortly. When $m = 0$, the energy equation $E = \sqrt{c^2\vec{p}^{\,2} + m^2c^4}$ becomes $E = c|\vec{p}|$; we can then rewrite these two equations as

$$\vec{\sigma}\cdot\hat{n}\,\phi_R = \phi_R, \tag{12.419}$$
$$\vec{\sigma}\cdot\hat{n}\,\phi_L = -\phi_L, \tag{12.420}$$

where $\hat{n} = \vec{p}/|\vec{p}|$ is the unit vector in the direction of the particle's momentum. These are known as the *decoupled Weyl equations*[48]. Since each one of the matrices σ_j is a 2×2 matrix, the states ϕ_R and ϕ_L must have two-components each; ϕ_R and ϕ_L are known as the *right-handed* and *left-handed Weyl, or chiral, spinors*, respectively (see Subsection 12.4.18 for a detailed treatment). So, the Weyl spinors are the two-component solutions of the Dirac equation for massless particles.

Next, since the helicity operator is given by $\widehat{\mathcal{H}} = \hat{\vec{S}}\cdot\hat{n} = (\hbar/2)\hat{\vec{\sigma}}\cdot\hat{n}$ as derived in Eq. (12.401), we obtain $\vec{\sigma}\cdot\hat{n} \equiv (2/\hbar)\widehat{\mathcal{H}}$. Inserting this relation into (12.419) and (12.420), we see that ϕ_R and ϕ_L *are helicity eigenstates with eigenvalues* $+\hbar/2$ and $-\hbar/2$, respectively:

$$\widehat{\mathcal{H}}\,\phi_R = \frac{\hbar}{2}\phi_R, \tag{12.421}$$
$$\widehat{\mathcal{H}}\,\phi_L = -\frac{\hbar}{2}\phi_L, \tag{12.422}$$

It should be noted that helicity is only a good quantum number for massless particles or massless antiparticles. From Eqs. (12.419) and (12.420), we can see that ϕ_R and ϕ_L are totally decoupled and independent of each other; they can thus be viewed as representing two separate particles. In fact, and as we will see in Subsection 12.4.17, *it has been experimentally determined that neutrinos are left-handed particles*[49], *while antineutrinos are right-handed*; hence, ϕ_L can be considered to describe spin-1/2 particles, whereas ϕ_R spin-1/2 antiparticles. Since ϕ_L and ϕ_R describe massless fermions and and antifermions, respectively, we can say that:
Massless fermions are purely left-handed and massless antifermions are purely right-handed.

Since a massless particle always moves at the speed of light, the value of the helicity is bound to be a Lorentz invariant.

From a historical perspective, Weyl was the first to have shown that the massless Dirac equation can be reduced to a much simpler equation that describes massless particles in terms

[48] As we will see below, since the Weyl equations (12.419)–(12.420) provide a reasonably accurate description of the properties of neutrinos, they are sometimes called the *neutrino equations*.

[49] It should be noted that the neutrinos we detect are produced (via weak interaction) in a left-handed state; this is usually attributed to the nature of weak interaction (due to its vector -axial structure).

of two-component spinors, as opposed to the four-component spinors that emerge from the Dirac equation.

In summary, Weyl provided us with a two-component theory and that the Dirac equation (12.708) is equivalent to the pair of Weyl equations or (12.412)–(12.413).

12.4.16 Weyl Spinors and Parity Violation

Let us examine the transformation of Weyl spinors under parity. As seen in Subsection 12.4.23, the parity operator $\hat{\mathcal{P}}$ transforms (or changes) vectors, $\vec{r} \longrightarrow -\vec{r}$ and $\vec{p} \longrightarrow -\vec{p}$, but leaves the spin unchanged, $\vec{\sigma} \longrightarrow \vec{\sigma}$, as it is an axial vector as illustrated in Eq. (12.528). Consequently, parity changes helicity: $\widehat{\mathcal{H}} \longrightarrow -\widehat{\mathcal{H}}$, since $\vec{\sigma}\cdot\vec{p} \longrightarrow -\vec{\sigma}\cdot\vec{p}$ or $\vec{\sigma}\cdot\hat{n} \longrightarrow -\vec{\sigma}\cdot\hat{n}$ and, hence, *parity transforms a right-handed spinor into a left-handed one and vice-versa*:

$$\hat{\mathcal{P}}\phi_R = \phi_L \qquad \text{and} \qquad \hat{\mathcal{P}}\phi_L = \phi_R. \tag{12.423}$$

Clearly, *Weyl spinors violate parity conservation*; therefore, *Weyl fermions do not obey spatial inversion*. Parity can thus be viewed as a test for *chirality* of a given phenomenon, since parity transforms this phenomenon into its mirror image. A chiral phenomenon is one that is not identical to its mirror image as will be seen in Subsection 12.4.18.

From (12.423), we can see that the Weyl spinors ϕ_R and ϕ_L transform into one another under parity action. So, taking one Weyl spinor alone, we cannot obtain parity conservation. However, taking both spinors together, we can construct a four-component spinor ψ that conserves parity

$$\phi_R \quad \text{and} \quad \phi_L \qquad \longrightarrow \qquad \psi = \begin{pmatrix} \phi_L \\ \phi_R \end{pmatrix}, \tag{12.424}$$

where ψ is nothing but the Dirac spinor. This shows how we can express the 4–component Dirac spinors ψ in terms of a combination of two 2–component Weyl spinors; this is achieved by simply stacking the left-handed Weyl spinor on top of the right-handed spinor[50]. So, stacking the left-handed Weyl spinor in the upper part of the Dirac spinor and the right handed Weyl spinor sits in the lower part, we have:

$$\phi_R = \begin{pmatrix} \phi_{R_1} \\ \phi_{R_2} \end{pmatrix} \quad \text{and} \quad \phi_L = \begin{pmatrix} \phi_{L_1} \\ \phi_{L_2} \end{pmatrix} \qquad \longrightarrow \qquad \psi = \begin{pmatrix} \phi_L \\ \phi_R \end{pmatrix} = \begin{pmatrix} \phi_{L_1} \\ \phi_{L_2} \\ \phi_{R_1} \\ \phi_{R_2} \end{pmatrix}. \tag{12.425}$$

So, although each one of the Weyl spinors violate parity conservation, the Dirac spinors do conserve parity. As a reminder, the Dirac spinors apply to massive fermions as opposed to the Weyl spinors which apply to massless fermions.

12.4.17 Weyl Fermions and Neutrinos

It is historically instructive to recall that when Weyl derived his equations in 1929, he had unwittingly stumbled upon the prediction of a new class of particles: *massless spin-1/2 fermions*

[50] See Subsection 12.4.18 for why stack ϕ_L on top of ϕ_R and not the other way around.

or *Weyl fermions*; these were not known to exist at that time, but they ended up playing an important role for decades after that.

The first hypothesis on the existence of a new spin-1/2 fermion with zero mass and zero charge was postulated in 1930 by Pauli in his "desperate" attempt to explain the apparent non-conservation of energy and momentum in radioactive decays. Then in 1933, Fermi proposed to name this new particle a *neutrino*[51] and used it to formulate his theory[52] of $\beta-$decay reactions which are mediated by weak interactions. The neutrino is conventionally denoted by the Greek letter ν; since every elementary particle has an antiparticle, the neutrino's antiparticle is the *antineutrino* which is denoted by $\overline{\nu}$. In 1989, CERN and SLAC researchers announced evidence that there are only three types of neutrinos: electron-neutrinos, ν_e, muon-neutrinos, ν_μ, and tau-neutrinos, ν_τ; their respective antineutrinos are $\overline{\nu}_e$, $\overline{\nu}_\mu$ and $\overline{\nu}_\tau$.

Neutrinos and antineutrinos are emitted in $\beta-$decay[53] reactions. We can find two types of $\beta-$decays: those that emit a β^+ particle (or positron) and a neutrino ν and those that emit a β^- particle (or electron) along with an antineutrino $\overline{\nu}$. In these radioactive decays, the total number of nucleons (neutrons + protons) remains the same in the nucleus in spite of the fact that the numbers of protons and neutrons in that atom's nucleus change. For instance, we can write, in general terms, the $\beta-$decay process of a radioactive element X (having A nucleons and Z protons) that decays into another element Y as follows

$$\beta^+ \text{ decay reaction:} \quad {}^{A}_{Z}X \longrightarrow {}^{A}_{Z-1}Y + {}_{+1}e^{+} + \nu \tag{12.426}$$

$$\beta^- \text{ decay reaction:} \quad {}^{A}_{Z}C \longrightarrow {}^{A}_{Z+1}Y + {}_{-1}e^{-} + \overline{\nu}. \tag{12.427}$$

Notice the conservation of A and Z in these decays. As an illustration, consider these two examples: a Carbon-10 isotope decays into Boron-10 by emitting a positron and an electron neutrino ν_e and, in another reaction, Carbon-14 decays into Nitrogen-14 by emitting an electron and an electron antineutrino $\overline{\nu}_e$:

$$\beta^+ \text{ decay reaction:} \quad {}^{10}_{6}C \longrightarrow {}^{10}_{5}B + {}_{+1}e^{+} + \nu_e \tag{12.428}$$

$$\beta^- \text{ decay reaction:} \quad {}^{14}_{6}C \longrightarrow {}^{14}_{7}N + {}_{-1}e^{-} + \overline{\nu}_e. \tag{12.429}$$

These equations show that in a β^+ decay, a proton in an atom's nucleus converts into a neutron, a positron, and a neutrino; and, on the other hand, during a β^- decay, a neutron in an atom's nucleus decays into a proton, an electron and an antineutrino:

$$\beta^+ \text{ decay of a proton:} \quad p \longrightarrow n + e^{+} + \nu_e \tag{12.430}$$

$$\beta^- \text{ decay of a neutron:} \quad n \longrightarrow p + e^{-} + \overline{\nu}_e. \tag{12.431}$$

We should note that the $\beta-$decay process enables unstable atoms to become more stable, for example, by transforming a neutron into a proton.

Wheyl fermions as candidate for neutrinos?

Since the properties of the neutrino (zero mass and zero charge) match those of the Weyl fermions, it was naturally assumed that the neutrinos are Weyl fermions. We should note

[51]The Italian word "neutrino" used to describe this particle is quite telling, because it is electrically neutral (*neutr*) and its rest mass is extremely small (*-ino*); the Italian equivalent of neutrino is the "little neutral one" or the "little neutron" as it resembles the neutron with no charge, a spin-1/2, and an extremely small mass or even massless.

[52]E. Fermi, *Nuovo Cimento*, **11**, 1–19 (1934).

[53]A β particle is either an electron (β^-) or a positron (β^+).

that the Weyl's theory faced initially (around 1930) what appeared to be a formidable flaw that, paradoxically, turned out to be one of its major strengths later on (1956): the Weyl spinors violate parity as (12.423) shows. The reason for this initial "false" alarm was due to the long held misconception that parity conservation is a universal property of nature and that all the fundamental interactions were thought to conserve parity. The ostensible universality of parity conservation was one of the most sacrosanct cornerstones of classical and modern physics; this is precisely why so many physicists shied away from questioning it. For instance, not accepting the violation of parity by Weyl fermions (neutrinos), Pauli had initially dismissed Weyl's theory. The interest in Weyl's work was then lost until 1956 when it was shown theoretically [54] and experimentally[55] that, unlike electromagnetic and strong interactions, the weak interaction does violate parity. The interest in the Weyl equations then recrudesced anew, particularly right after 1956 when parity violation has redemptively given solid credibility to Weyl's analysis[56].

On a related note, ever since neutrinos were postulated in 1930, they remained elusive until they were detected for the first time in 1956 by Clyde Cowan and Fred Reines[57] using a fission reactor as a source of neutrinos. Additionally, the 1957 Goldhaber-Experiment[58], which was aimed at determining the helicity of neutrinos, has confirmed the following property:

All neutrinos are left-handed and all antineutrinos are right-handed.

That is, *all neutrinos have their spin vectors opposite to their momentum vectors (i.e., left-handed), while all antineutrinos have spin parallel to momentum (i,e., right-handed);* the eigenvalues of the helicity operator are -1 for ν and $+1$ for $\overline{\nu}$. So, neutrinos are described by the left-handed Weyl spinors ϕ_L. This causes *the weak interactions that emit neutrinos and antineutrinos to violate parity conservation.* This experimental finding is compatible with Weyl's theoretical result of Eq. (12.423). We may also mention that there is no experimental evidence for right-handed neutrinos nor for left-handed antineutrinos.

The time-line of the above mentioned experimental findings can be summarized as follows,

- The solutions of Weyl's equations are massless fermions which violate parity (1929),
- Neutrinos (massless and chargeless fermions) were postulated in 1930 and first detected in 1956,
- Neutrinos were confirmed to violate parity in 1957,
- All neutrinos are left-handed and all antineutrinos are right-handed (1957),
- Neutrinos were thought to be massless for a long time.

---

[54] T. D. Lee and C. N. Yang, *Phys. Rev.*, **104**, 254 (1956); T. D. Lee and C. N. Yang, *Phys. Rev.*, **105**, 1671 (1957); A. Salam, *Nuovo Cimento*, **5**, 299 (1957).

[55] C. S. Wu, E. Amblerand, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105**, 1413 (1957); R. Garwin, L. Lederman, M. Weinrich, Phys. Rev. **105**, 1415 (1957).

[56] While Weyl died in 1955 and, hence, had missed the chance to see the 1956 triumph of his theory, Pauli lived long enough to witness the failure of his long-held view of parity conservation in $\beta-$decay reactions (Pauli died in 1958).

[57] C. L. Cowan, Jr., F. Reines, F. B. Harrison, H. W. Kruse and A. D. McGuire, *Detection of the Free Neutrino: A Confirmation*, Science **124**, 103 (1956).
Frederick Reines and Clyde L. Cowan, Jr., *The Neutrino*, Nature **178**, 446 (1956).

[58] M. Goldhaber, L. Grodzins, and A. W. Sunyar, *Phys. Rev.* **109**, 1015 (1958).

All these findings are compatible with the properties of the Weyl's spinors. As a result, and for decades, *neutrinos were thought to be Weyl fermions;* neutrinos were described by ϕ_L and antineutrinos by ϕ_R.

However, in 1998 when it was confirmed experimentally[59] that *neutrinos have very small, but nonzero, masses, it became evident that they cannot be Weyl fermions.* But the search for them never stopped. It is only recently (2015) that Weyl fermions were detected[60]. In this experiment, Weyl fermions were realized as emergent quasiparticles in a low-energy condensed matter system. This has resulted in introducing the concept of a Weyl semimetal which is considered to be a solid state crystal whose low energy excitations are Weyl fermions that carry electrical charge even at room temperatures[61].

Additionally, in a 2017 experiment[62], it was reported that "Weyl fermions" were discovered in materials with strong interaction between electrons and that they only exist as "quasiparticles" within a solid material. It was also found that, despite having no mass, these Weyl fermions move very slowly. Clearly, since their introduction in 1929, the Weyl fermions continue to attract attention and, with these new discoveries, they have the potential to be very relevant in new technological applications.

Dirac and Weyl fermions

Up to now, we have seen two types of spin-1/2 particles: one proposed by Dirac and the other by Weyl. On the one hand, the Dirac fermions are spin-1/2 particles with nonzero mass that emerge naturally from the *massive Dirac equation* and are represented by four-component complex spinors describing relativistic electrons. On the other hand, we have the Weyl fermions that are spin-1/2 *massless* particles that are described by two-component spinors that are solutions of the *massless Dirac equation* (i.e., the Weyl equations); their helicity is a good quantum number, since the Weyl spinors describe particles that move at the speed of light.

12.4.18 Chirality

An object or a system (e.g., a molecule) is said to be *chiral* if it is not superimposable on its mirror image. Hence, chirality[63] is a property of asymmetry, for it pertains to the handedness of the object in being either left-chiral or right-chiral.

In this subsection, we want to discuss the chiral symmetry of the Dirac spinors. As seen above, the Weyl spinors (corresponding to massless particles) are either left-handed or right-handed.

Chirality, also called handedness, is associated with the matrix γ^5 which was introduced in Eq. (12.188); γ^5 is the chirality operator and its eigenvalue is called chirality, as will be

[59] Y. Fukuda *et al.* (Super-Kamiokande Collaboration), *Phys. Rev. Lett.*, **81**, 1562 (1998); S. Hatakeyama *et al.* (Kamiokande Collaboration) *Phys. Rev. Lett.* **81**, 2016 (1998); C. Athanassopoulos *et al.* (LSND Collaboration) *Phys. Rev. Lett.* **81**, 1774 (1998).

[60] D. Ciudad, *Massless yet real*, Nature Mater **14**, 863 (2015).
Hamish Johnston, *Weyl fermions are spotted at long last*, Physics World (23 July 2015).

[61] X. Wan, A. M. Turner, A. Vishwanath, and S. Y. Savrasov, *Topological Semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates*, Phys. Rev. B. **83 (20)**, 205101 (2011).
A. A. Burkov and L. Balents, *Weyl Semimetal in a Topological Insulator Multilayer*, Phys. Rev. Lett. **107 (12)**, 127205 (2011).

[62] Hsin-Hua Lai, Sarah E. Grefe, Silke Paschen, and Qimiao Si, *"Weyl–Kondo semimetal in heavy-fermion systems"*. *Proceedings of the National Academy of Sciences*, 2017.

[63] "Chirality" has its origin in the Greek word "hand."

seen below. For massless fermions, and as shown in Eqs. (12.412)–(12.413), the Dirac equation reduces to the Weyl equations

$$i\gamma^{\mu}\partial_{\mu}\psi = 0. \tag{12.432}$$

It is easy to verify that $\gamma^5\psi$ is also solution to (12.432):

$$i\gamma^{\mu}\partial_{\mu}\left(\gamma^5\psi\right) = 0. \tag{12.433}$$

Next, we can use γ^5 to express any Dirac spinor ψ in terms of a right-chiral and left-chiral part as follows:

$$\psi = \left[\frac{1}{2}(I_4 + \gamma^5) + \frac{1}{2}(I_4 - \gamma^5)\right]\psi = \psi_R + \psi_L, \tag{12.434}$$

where ψ_R and ψ_L are respectively known as the right-handed and left-handed chiral projections of ψ,

$$\psi_R = \frac{1}{2}(I_4 + \gamma^5)\psi \equiv \hat{P}_R\psi, \qquad \psi_L = \frac{1}{2}(I_4 - \gamma^5)\psi \equiv \hat{P}_L\psi, \tag{12.435}$$

and where $\hat{P}_R$ and $\hat{P}_L$ are the *chirality projection operators*:

$$\hat{P}_R = \frac{1}{2}(I_4 + \gamma^5), \qquad \hat{P}_L = \frac{1}{2}(I_4 - \gamma^5). \tag{12.436}$$

As shown in Eq. (12.192), γ^5 is a projection operator, since it is both Hermitian and equal to its square. Thus, we can easily show that $\hat{P}_R$ and $\hat{P}_L$ are also projection operators since both of them are Hermitian and equal to their respective squares:

$$\hat{P}_R^{\dagger} = \hat{P}_R, \quad \hat{P}_R^2 = \hat{P}_R, \qquad \hat{P}_L^{\dagger} = \hat{P}_L, \quad \hat{P}_L^2 = \hat{P}_L. \tag{12.437}$$

As mentioned at the end of Subsection 12.4.4, when we deal with ultra-relativistic phenomena, it is more convenient to use the Weyl representation of the *gamma*–matrices. For instance, using the Weyl basis (12.226) where γ^5 is given by

$$\gamma^5 = \begin{pmatrix} -I_2 & 0 \\ 0 & I_2 \end{pmatrix} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{12.438}$$

we can easily show that $\hat{P}_R$ and $\hat{P}_L$ are 4 matrices. For this, we need only to insert (12.438) into (12.436):

$$\hat{P}_R = \frac{1}{2}(I_4 + \gamma^5) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \hat{P}_L = \frac{1}{2}(I_4 - \gamma^5) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{12.439}$$

We can verify that these operators satisfy the following properties:

$$\hat{P}_R + \hat{P}_L = I_4, \qquad \hat{P}_R^2 = \hat{P}_R, \qquad \hat{P}_L^2 = \hat{P}_L, \qquad \hat{P}_R\hat{P}_L = \hat{P}_L\hat{P}_R = 0. \tag{12.440}$$

Since $\hat{P}_R$ and $\hat{P}_L$ are 4×4 matrices, the chiral states ψ_R and ψ_L are also 4–component spinors. We can express them in terms of the Weyl spinors ϕ_R and ϕ_L, that were introduced in Eq. (12.425), as follows

$$\psi_R = \begin{pmatrix} 0 \\ \phi_R \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 0 \\ \phi_{R_1} \\ \phi_{R_2} \end{pmatrix}, \qquad \psi_L = \begin{pmatrix} \phi_L \\ 0 \end{pmatrix} \equiv \begin{pmatrix} \phi_{L_1} \\ \phi_{L_2} \\ 0 \\ 0 \end{pmatrix} \longrightarrow \psi = \begin{pmatrix} \phi_L \\ \phi_R \end{pmatrix} = \begin{pmatrix} \phi_{L_1} \\ \phi_{L_2} \\ \phi_{R_1} \\ \phi_{R_2} \end{pmatrix}, \tag{12.441}$$

where ϕ_R and ϕ_L are two-component spinors. If we apply the 4×4 matrices (12.439) on the spinors (12.441), we would recover the relations $\hat{P}_R\psi_R = \psi_R$ and $\hat{P}_L\psi_R = \psi_L$:

$$\hat{P}_R\psi = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \phi_{L_1} \\ \phi_{L_2} \\ \phi_{R_1} \\ \phi_{R_2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \phi_{R_1} \\ \phi_{R_2} \end{pmatrix} \equiv \psi_R, \qquad \hat{P}_L\psi = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \phi_{L_1} \\ \phi_{L_2} \\ \phi_{R_1} \\ \phi_{R_2} \end{pmatrix} = \begin{pmatrix} \phi_{L_1} \\ \phi_{L_2} \\ 0 \\ 0 \end{pmatrix} \equiv \psi_L. \tag{12.442}$$

Note that when the chirality projection operators act on states with different handedness, they yield zero projections:

$$\hat{P}_R\psi_L = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \phi_{L_1} \\ \phi_{L_2} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \hat{P}_L\psi_R = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \phi_{R_1} \\ \phi_{R_2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{12.443}$$

Next, using the fact that $(\gamma^5)^2 = I_4$, we can verify that ψ_R and ψ_L are eigenstates of γ^5 with eigenvalues $+1$ and -1, respectively:

$$\gamma^5\psi_R = \frac{1}{2}\left[\gamma^5 + (\gamma^5)^2\right]\psi = \frac{1}{2}(1 + \gamma^5)\psi = \psi_R, \tag{12.444}$$

$$\gamma^5\psi_L = \frac{1}{2}\left[\gamma^5 - (\gamma^5)^2\right]\psi = -\frac{1}{2}(1 - \gamma^5)\psi = -\psi_L. \tag{12.445}$$

Since $\gamma_5^2 = I_4$, the eigenvalues of γ_5 are ± 1 (see (12.192). The eigenvectors of γ_5 corresponding the $+1$ eigenvalue are said to have right-handed or positive chirality and those to the -1 eigenvalue are left-handed or negative chirality; the states ψ_R and ψ_L displayed in (12.444)–(12.445) have positive and negative chiralities, respectively. So, the chirality of a particle is determined by whether its spinor transforms in a right-handed or left-handed representation.

Remarks

- Chirality is not conserved, because γ^5 does not commute with the Dirac Hamiltonian even for a free particle at rest. This is due to the fact that γ^5 does not commute with the mass term $\gamma^0 mc^2$ in the Hamiltonian (12.228). This can be seen as follows: while γ^5 commutes with the kinetic energy term $c\gamma^0\gamma^j\hat{p}^j$ (where $j = 1, 2, 3$), it anticommutes with the mass term $\gamma^0 mc^2$; that is, since γ^5 anticommutes with γ^μ (i.e., $\{\gamma^5, \gamma^\mu\} = 0$ where $\mu = 0, 1, 2, 3$), and since the kinetic energy term has two γ matrices while the mass term has only only one, we have

$$\left[\gamma^0\gamma^j, \gamma^5\right] = 0, \qquad \left[\gamma^0, \gamma^5\right] = 2\gamma^0\gamma^5 \quad \Longrightarrow \quad \left[\hat{H}, \gamma^5\right] \neq 0, \tag{12.446}$$

and hence

$$\left[\hat{H},\ \gamma^5\right] = \left[\gamma^0\vec{\gamma}\cdot\hat{\vec{p}} + \gamma^0 mc^2,\ \gamma^5\right] = c\left[\gamma^0\gamma^i\hat{p}^i,\ \gamma^5\right] + mc^2\left[\gamma^0,\ \gamma^5\right] = 2mc^2\gamma^0\gamma^5. \quad (12.447)$$

- From (12.447), we see that if the mass is zero, chirality is conserved, since the Hamiltonian commutes with γ^5:

$$\left[\hat{H},\ \gamma^5\right] = 2mc^2\gamma^0\gamma^5 = 0. \quad (12.448)$$

- As shown in Problem 12.32 on Page 874, the parity operator transforms a left-chiral state into a right-chiral state and vice versa: $\hat{\mathcal{P}}\psi_R = \pm\psi_L$ and $\hat{\mathcal{P}}\psi_L = \pm\psi_R$.

Differences between helicity and chirality

There are a lot of subtleties about the connection between helicity and chirality. To clear some of these misconceptions, let us look at the differences and similarities between them. From the very outset, we need to state that helicity and chirality are two different quantities. Helicity is the projection of the spin along the direction of motion, whereas chirality is simply an asymmetry characteristic of objects or phenomena. While the helicity operator $\widehat{\mathcal{H}} = \hat{\vec{S}}\cdot\left(\hat{\vec{p}}\,/\,|\vec{p}|\right)$ has the physical dimensions of an angular momentum, the chirality operator γ^5 has no physical dimensions, for it simply is a dimensionless pseudoscalar. Contrary to chirality which is an inherent or intrinsic property of a particle, helicity is an extrinsic physical property resulting from the alignment or anti-alignment of the particle's spin and momentum. Hence, we see that the helicity is frame-dependent, while chirality is frame-independent, since the helicity is determined by frame-dependent quantities (momentum) while the chirality does not.

However, in spite of being different quantities, helicity and chirality do display some similar characteristics, especially in the ultra relativistic domain; i.e., when $v \longrightarrow c$ or $m = 0$ as we will see shortly. So, let us consider the two cases of massive and massless particles.

First, for massive particles, while the helicity of a free particle is conserved in time, chirality is not; this is self evident from (12.447) and the fact that the helicity operator commutes with the Hamiltonian (see Eq. (12.407)):

$$\left[\hat{H},\ \gamma^5\right] = 2mc^2\gamma^0\gamma^5, \qquad [\hat{H},\ \widehat{\mathcal{H}}] = 0. \quad (12.449)$$

Additionally, chirality is Lorentz invariant, but helicity is not. This can be easily inferred from the fact that while chirality commutes with the generator of the Lorentz transformation $\sigma^{\mu\nu}$, helicity does not; using (12.186), we can show that

$$\left[\sigma^{\mu\nu},\ \gamma^5\right] \equiv \frac{i}{2}\left[[\gamma^\mu,\ \gamma^\nu],\ \gamma^5\right] = 0, \qquad \left[\sigma^{\mu\nu},\ \widehat{\mathcal{H}}\right] \neq 0. \quad (12.450)$$

The first commutator is zero because γ^5 commutes with an even product of γ−matrices and the second commutator is not zero because the helicity operator does not commute with γ−matrices. So, helicity is not Lorentz invariant in the case of a massive particle.

Second, for the *special case of a massless particle*, helicity and chirality display similar properties. On the one hand, both of the helicity and the chirality are conserved, since they both commute with the Hamiltonian as can be seen from (12.407) and (12.448):

$$[\hat{H},\ \widehat{\mathcal{H}}] = 0, \qquad , \qquad \left[\hat{H},\ \gamma^5\right] = 2mc^2\gamma^0\gamma^5 = 0. \quad (12.451)$$

This also evident from Eqs. (12.421)–(12.422) and (12.444)–(12.445), which show that, when $m = 0$, both of the helicity and the chirality operators preserve the chiralities of the Dirac spinors[64]:

$$\widehat{\mathcal{H}}\,\phi_R = \frac{\hbar}{2}\phi_R, \qquad \widehat{\mathcal{H}}\,\phi_L = -\frac{\hbar}{2}\phi_L, \tag{12.452}$$

$$\gamma^5\psi_R = \psi_R, \qquad \gamma^5\psi_L = -\psi_L. \tag{12.453}$$

So, in the massless limit, the Dirac equation shows that a particle of positive helicity has positive chirality, and vice versa. On the other hand, the helicity of massless particles is a relativistic invariant and it always matches the massless particles' chirality; helicity is not Lorentz invariant unless the particle is massless. If the particle has mass, one can always make a Lorentz transformation into an inertial frame with a speed larger than the particle's and, hence, be able to invert or flip the helicity. If, however, the particles travels at the speed of light (i.e., $m = 0$), there is no inertial frame that travels faster than light and,hence, helicity becomes a Lorentz invariant.

In short, in the case of massive particles, helicity is conserved but not a Lorentz invariant, whereas chirality is not conserved but is a Lorentz invariant; hence, helicity and chirality are not the same for massive particles. However, in the case of massless particles, the helicity and the chirality are both conserved and Lorentz invariant; clearly, in the limit of $E \gg mc^2$ or massless limit, there is a one-to-one correspondence between chirality and helicity.

12.4.19 Dirac Equation in an Electromagnetic Field

Consider a relativistic particle of charge q (where $q = -e$ for an electron) which is moving under the influence of an electromagnetic field $A^\mu = (U/c, \vec{A})$ where U is the electric scalar potential and $\vec{A}$ is the vector potential. In this subsection, we want to derive the Dirac equation for such a particle.

Substituting (12.123) and (12.124) into the free particle Dirac equation (12.150), we obtain

$$i\hbar\frac{\partial\psi}{\partial t} = \left(c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2\right)\psi \quad \longrightarrow \quad i\hbar\frac{\partial\psi}{\partial t} = \left[c\vec{\alpha}\cdot\left(\hat{\vec{p}} - q\vec{A}\right) + \beta mc^2 + \hat{V}\right]\psi, \tag{12.454}$$

where $\hat{V}$ is defined by[65] $\hat{V}(\vec{r}) = qU(\vec{r})$; it represents the interaction between the charge q and the electric scalar potential U. For simplicity, we are considering the case where there is no external potential $\hat{V}(\vec{r})$; hence, the charge interacts only with the external electromagnetic field. However, if, in addition to interacting with an external electromagnetic field, the particle is also subjected to an external potential $\hat{V}(\vec{r})$, the formalism below can handle it quite readily; all we need to do is to simply replace $qU(\vec{r})$ by $(qU(\vec{r}) + \hat{V}(\vec{r}))$ in all relevant expressions below.

Relation (12.454) is known as the Dirac equation for a charge particle q in an electromagnetic field $A^\mu = (U/c, \vec{A})$. Similarly, inserting (12.123) into the Dirac Hamiltonian (12.147), we obtain

$$\hat{H} = c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2 \quad \longrightarrow \quad \hat{H} = c\vec{\alpha}\cdot\left(\hat{\vec{p}} - q\vec{A}\right) + \beta mc^2 + \hat{V}, \tag{12.455}$$

[64]The factor $\hbar/2$ is due to the fact that, as mentioned above, the helicity has the physical dimensions of an angular momentum, whereas γ^5 is dimensionless.

[65]One needs to be careful not to confuse the $\hat{V}(\vec{r})$ in $\hat{V}(\vec{r}) = qU(\vec{r})$ with the external potential energy which is taken to be zero here because the Dirac equation (12.150) is for a free particle.

which, using (12.161), becomes

$$\hat{H} = \begin{pmatrix} \left(mc^2 + \hat{V}\right) I_2 & c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right) \\ c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right) & \left(-mc^2 + \hat{V}\right) I_2 \end{pmatrix} \equiv \begin{pmatrix} mc^2 + \hat{V} & c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right) \\ c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right) & -mc^2 + \hat{V} \end{pmatrix}, \tag{12.456}$$

where I_2 is the 2×2 unit matrix and the Pauli matrices $\vec{\sigma}$ are given by (12.162). Now, by expressing the 4–component Dirac spinor ψ in terms of two 2–component spinors ϕ_A and ϕ_B as $\psi = \begin{pmatrix} \phi_A \\ \phi_B \end{pmatrix}$ and applying it to (12.456),

$$\begin{pmatrix} mc^2 + \hat{V} & c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right) \\ c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right) & -mc^2 + \hat{V} \end{pmatrix} \begin{pmatrix} \phi_A \\ \phi_B \end{pmatrix} = \hat{H} \begin{pmatrix} \phi_A \\ \phi_B \end{pmatrix}, \tag{12.457}$$

we obtain the following two coupled equations

$$\left(mc^2 + \hat{V}\right)\phi_A + c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right)\phi_B = E\phi_A \tag{12.458}$$

$$c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right)\phi_A + \left(-mc^2 + \hat{V}\right)\phi_B = E\phi_B, \tag{12.459}$$

or

$$\left(E - mc^2 - \hat{V}\right)\phi_A = c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right)\phi_B \tag{12.460}$$

$$\left(E + mc^2 - \hat{V}\right)\phi_B = c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right)\phi_A. \tag{12.461}$$

In the next subsection, we are going to use these two equations to study the non-relativistic limit of the Dirac equation.

12.4.20 Non-relativistic limit of the Dirac Equation

While non-relativistic quantum mechanics is founded on the Schrödinger equation, relativistic quantum mechanics is based on the Dirac equation. The multifaceted successes of the Schrödinger equation in the non-relativistic domain are unassailable. The key question now is to find if the Dirac equation has the correct non-relativistic limit; one of the most important tests of the Dirac equation is its ability to give us back the Schrödinger equation in the non-relativistic limit. For this, we are going to expand the Hamiltonian in terms of the small parameter $1/c$ in two steps, first up to order $1/c$ and then up to $1/c^2$.

In the non-relativistic limit, the particle's rest-mass energy mc^2 is the dominant term in the energy expression $E = E_{NR} + mc^2$, where E_{NR} is the energy of the particle (electron) in addition to its rest-mass energy (this can be the particle's kinetic energy, the potential energy and any other energy term); in short, $E_{NR} = E - mc^2$ is the nonrelativistic part of the particle's energy with $E_{NR} \ll mc^2$. Inserting $E = E_{NR} + mc^2$ into (12.461) and rearranging some terms, we obtain

$$\begin{aligned} \phi_B &= \frac{1}{E_{NR} + 2mc^2 - \hat{V}} c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right)\phi_A \\ &= \frac{1}{2mc^2}\left(1 + \frac{E_{NR} - \hat{V}}{2mc^2}\right)^{-1} c\vec{\sigma} \cdot \left(\hat{\vec{p}} - q\vec{A}\right)\phi_A. \end{aligned} \tag{12.462}$$

12.4.20.1 First-Order Approximation of the Dirac Equation

In the non-relativistic limit, we have $(E_{NR} - \hat{V}) \ll mc^2$. Hence, an expansion of (12.462) in terms of the small parameter $(E_{NR} - \hat{V})/(2mc^2)$ *up to the first-order* in $1/c$ gives

$$\phi_B \simeq \frac{1}{2mc}\vec{\sigma}\cdot\left(\hat{\vec{p}} - q\vec{A}\right)\phi_A, \tag{12.463}$$

which, when inserted into the right-hand side of (12.460), leads to

$$\left(E_{NR} - \hat{V}\right)\phi_A \simeq \frac{1}{2m}\left[\vec{\sigma}\cdot\left(\hat{\vec{p}} - q\vec{A}\right)\right]^2\phi_A \tag{12.464}$$

or to

$$\left\{\frac{1}{2m}\left[\vec{\sigma}\cdot\left(\hat{\vec{p}} - q\vec{A}\right)\right]^2 + \hat{V}\right\}\phi_A = E_{NR}\phi_A. \tag{12.465}$$

Now, let us focus on the calculation of $\left[\vec{\sigma}\cdot\left(\hat{\vec{p}} - q\vec{A}\right)\right]^2$. Using the relation $\sigma_j\sigma_k = \delta_{jk}I_2 + i\varepsilon_{jkl}\sigma_l$ from Chapter 5, we can verify that, for any two arbitrary vectors $\vec{C}$ and $\vec{D}$ which commute with $\vec{\sigma}$, we have

$$(\vec{\sigma}\cdot\vec{C})(\vec{\sigma}\cdot\vec{D}) = (\vec{C}\cdot\vec{D})I_2 + i\vec{\sigma}\cdot(\vec{C}\times\vec{D}). \tag{12.466}$$

Using this relation, we can write

$$\left[\vec{\sigma}\cdot\left(\hat{\vec{p}} - q\vec{A}\right)\right]\left[\vec{\sigma}\cdot\left(\hat{\vec{p}} - q\vec{A}\right)\right] = \left(\hat{\vec{p}} - q\vec{A}\right)^2 + i\vec{\sigma}\cdot\left[\left(\hat{\vec{p}} - q\vec{A}\right)\times\left(\hat{\vec{p}} - q\vec{A}\right)\right], \tag{12.467}$$

where, using $\hat{\vec{p}}\times\hat{\vec{p}} = 0$ and $\vec{A}\times\vec{A} = 0$, we can write the last term as

$$\left(\hat{\vec{p}} - q\vec{A}\right)\times\left(\hat{\vec{p}} - q\vec{A}\right) = -q\left(\hat{\vec{p}}\times\vec{A} + \vec{A}\times\hat{\vec{p}}\right) + \hat{\vec{p}}\times\hat{\vec{p}} + q^2\vec{A}\times\vec{A} \equiv -q\left(\hat{\vec{p}}\times\vec{A} + \vec{A}\times\hat{\vec{p}}\right). \tag{12.468}$$

Next, if we use $\hat{\vec{p}} = -i\hbar\vec{\nabla}$, we can rewrite the last term as

$$\hat{\vec{p}}\times\vec{A} + \vec{A}\times\hat{\vec{p}} = -i\hbar\left(\vec{\nabla}\times\vec{A} + \vec{A}\times\vec{\nabla}\right). \tag{12.469}$$

Now, using the identity[66] $\vec{\nabla}\times(\vec{A}\phi) = \left(\vec{\nabla}\times\vec{A}\right)\phi - \vec{A}\times(\vec{\nabla}\phi)$ along with the relation $\vec{B} = \vec{\nabla}\times\vec{A}$ which relates the magnetic field $\vec{B}$ to the vector potential $\vec{A}$, we can recast (12.469) in the following form:

$$\vec{\nabla}\times\vec{A}\phi + \vec{A}\times\vec{\nabla}\phi = \left(\vec{\nabla}\times\vec{A}\right)\phi - \vec{A}\times(\vec{\nabla}\phi) + \vec{A}\times(\vec{\nabla}\phi) = \left(\vec{\nabla}\times\vec{A}\right)\phi \equiv \vec{B}\phi. \tag{12.470}$$

Now, plugging (12.470) into (12.469), we obtain

$$\hat{\vec{p}}\times\vec{A} + \vec{A}\times\hat{\vec{p}} = -i\hbar\vec{B}, \tag{12.471}$$

which, in turn, when inserted into the last term of (12.468) leads to

$$\left(\hat{\vec{p}} - q\vec{A}\right)\times\left(\hat{\vec{p}} - q\vec{A}\right) = -q\left(\hat{\vec{p}}\times\vec{A} + \vec{A}\times\hat{\vec{p}}\right) = iq\hbar\left(\vec{\nabla}\times\vec{A}\right) = iq\hbar\vec{B}, \tag{12.472}$$

[66] In this identity, the operator $\vec{\nabla}$ in the "$\left(\vec{\nabla}\times\vec{A}\right)\phi$" term acts only on $\vec{A}$, not on ϕ nor on anything else on the right-hand side.

Finally, plugging (12.472) into (12.467), we obtain

$$\left[\vec{\sigma}\cdot\left(\hat{\vec{p}}-q\vec{A}\right)\right]\left[\vec{\sigma}\cdot\left(\hat{\vec{p}}-q\vec{A}\right)\right]=\left(\hat{\vec{p}}-q\vec{A}\right)^2-q\hbar\vec{\sigma}\cdot\vec{B}, \tag{12.473}$$

which, in turn, when substituted into (12.465), leads to

$$\left[\frac{1}{2m}\left(\hat{\vec{p}}-q\vec{A}\right)^2+\hat{V}-\frac{q\hbar}{2m}\vec{\sigma}\cdot\vec{B}\right]\phi_A=E_{NR}\phi_A, \tag{12.474}$$

which, when combined with $\hat{\vec{S}}=\frac{\hbar}{2}\vec{\sigma}$, reduces to

$$\boxed{\left[\frac{1}{2m}\left(\hat{\vec{p}}-q\vec{A}\right)^2+\hat{V}-\frac{q}{m}\hat{\vec{S}}\cdot\vec{B}\right]\phi_A=E_{NR}\phi_A.} \tag{12.475}$$

This is the familiar *Pauli equation*[67] which was derived by Pauli from the Schrödinger equation for a spin 1/2 particle of charge q in an electromagnetic field, where ϕ_A is the (2−component) Pauli spinor; this equation is also called the *Schrödinger-Pauli equation* (see Subsection 6.4 on Page 403). This result shows that the Dirac equation reduces to the Schrödinger equation in the nonrelativistic limit.

The terms $\frac{1}{2m}\left(\hat{\vec{p}}-q\vec{A}\right)^2$ and $\hat{V}$ in (12.475) are the usual kinetic and potential energy terms of the Schrödinger equation; the two components of the Pauli spinor ϕ_A represent the internal spin of the electron. While the term $-\frac{q\hbar}{2m}\vec{\sigma}\cdot\vec{B}$ was added by Pauli to the Schrödinger equation in a rather ad hoc manner, it appears naturally within the Dirac equation; this term represents the anomalous Zeeman interaction energy of the particle's magnetic dipole moment $\vec{\mu}_S$ with an external magnetic field $\vec{B}$:

$$\hat{H}_S=-\frac{q\hbar}{2m}\vec{\sigma}\cdot\vec{B}\equiv-\hat{\vec{\mu}}_S\cdot\vec{B}\quad\Longrightarrow\quad\hat{\vec{\mu}}_S=\frac{q\hbar}{2m}\vec{\sigma}\equiv\frac{q}{m}\hat{\vec{S}}, \tag{12.476}$$

where $\hat{\vec{\mu}}_S$ is known as the intrinsic or *spin magnetic moment* of the spin 1/2 particle.

As seen in Chapter 5 (see Eq. (5.91) on Page 327), the classical orbital magnetic moment is given by

$$\hat{\vec{\mu}}_L=\frac{q}{2m}\hat{\vec{L}}=\gamma_L\hat{\vec{L}}, \tag{12.477}$$

where $\gamma_L=q/(2m)$ is the orbital *gyromagnetic ratio*. A simple comparison of (12.476) and (12.477) shows that the spin magnetic moment of a spin 1/2 particle is *twice* that obtained from classical physics. As such, we can express $\hat{\vec{\mu}}_S$ in terms of *gyromagnetic factor* $g_s=2$ as follows

$$\boxed{\hat{\vec{\mu}}_S=\frac{q}{m}\hat{\vec{S}}=g_s\frac{q}{2m}\hat{\vec{S}}\equiv\gamma_s\hat{\vec{S}}\quad\Longrightarrow\quad\text{gyromagnetic ratio}\quad g_s=2,} \tag{12.478}$$

where $\gamma_s=q/m$ is the *spin gyromagnetic ratio* The $g_s=2$ value is one of the most important triumphs of the Dirac equation, for it agrees with the experimental value to an outstanding accuracy[68], $g_{EXP}=2.0023193\cdots$.

[67] W. Pauli, *Z. Phys.* **43**, 601 (1927).

[68] H. M. Foley and P. Kusch, *Phys. Rev.* **73**, 412 (1948).

We may mention the experimental values of the gyromagnetic factors for free nucleons: $g_p = 5.586$ for a proton and $g_N = -3.826$ for a neutron. These values are due to the fact that protons and neutrons are not point-like particles; they have internal structure and each is made of three quarks. It is customary to express the proton's gyromagnetic ratio γ_p in terms of the *nuclear magneton* μ_N as follows:

$$\gamma_p = \frac{g_p \mu_N}{\hbar}, \qquad \mu_N = \frac{e\hbar}{2m_p} = 5.0507 \times 10^{-27}\ \text{J/T}, \qquad g_p = 5.586, \tag{12.479}$$

where e and m_p are the charge and mass of the proton, respectively.

12.4.20.2 Second-Order Approximation of the Dirac Equation

If the calculations involved in the nonrelativistic limit of the Dirac equation are pushed up to the *second order* in $1/c^2$, we can show that the well known correction terms, such as the relativistic correction to the kinetic energy, the spin-orbit coupling term and the Darwin term emerge naturally from the Dirac formalism. However, in nonrelativistic quantum mechanics, these terms are introduced in an *ad hoc* manner; contributions from these terms were calculated in Chapter 9 using perturbation theory (see Subsection 9.2.3 on Page 548).

Next, to carry out the second-order calculations, we are going to take $\vec{A} = 0$ (for simplicity) and expand (12.462) up to order $1/c^2$,

$$\phi_B = \frac{1}{2mc^2}\left(1 + \frac{E_{NR} - \hat{V}}{2mc^2}\right)^{-1} c\vec{\sigma}\cdot\hat{\vec{p}}\,\phi_A \simeq \frac{1}{2mc}\left(1 - \frac{E_{NR} - \hat{V}}{2mc^2}\right)\vec{\sigma}\cdot\hat{\vec{p}}\,\phi_A. \tag{12.480}$$

If we insert this expression into the right-hand side of (12.460), and carry out the ensuing calculations (see Problem 12.33 on Page 875), we end up with the following non-relativistic limit of the Dirac Hamiltonian

$$\hat{H}_{NR} = \frac{\hat{\vec{p}}^{\,2}}{2m} + \hat{V} - \frac{\hat{\vec{p}}^{\,4}}{8m^3c^2} + \frac{1}{2m^2c^2}\hat{\vec{S}}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right) + \frac{q\hbar^2}{8m^2c^2}(\nabla^2\hat{V}) \tag{12.481}$$

which can be rewritten as

$$\boxed{\hat{H}_{NR} = \frac{\hat{\vec{p}}^{\,2}}{2m} + \hat{V} + \hat{H}_R + \hat{H}_{SO} + \hat{H}_{Darwin},} \tag{12.482}$$

where

$$\boxed{\hat{H}_R = -\frac{\hat{\vec{p}}^{\,4}}{8m^3c^2}, \qquad \hat{H}_{SO} = \frac{1}{2m^2c^2}\hat{\vec{S}}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right), \qquad \hat{H}_{Darwin} = \frac{\hbar^2}{8m^2c^2}(\nabla^2\hat{V}).} \tag{12.483}$$

The first two terms in (12.482) are the familiar kinetic and potential energy terms.

The third term, $\hat{H}_R$, is the relativistic correction to the kinetic energy; it results from the relativistic variation of the mass with speed[69].

The fourth term, $\hat{H}_{SO}$, is the spin-orbit coupling term; it gives the energy of interaction between the magnetic moment $\vec{\mu}_S = (q/m)\hat{\vec{S}}$ of the electron and the effective magnetic field $\vec{B}$

[69]The standard relativistic mass expression $m = m_0/\sqrt{1 - v^2/c^2}$ shows that m increases as the speed v increases.

due to the electric current created by the orbital motion of the electron around the nucleus. In the case of a central (Coulomb) potential $\hat{V} \propto 1/r$ and since $\vec{\nabla}\hat{V} = \frac{\vec{r}}{r}\frac{d\hat{V}}{dr}$, we can write the spin-orbit term as

$$\hat{H}_{SO} = \frac{1}{2m^2c^2}\hat{\vec{S}}\cdot\left(\frac{d\hat{V}}{dr}\frac{\vec{r}}{r}\times\hat{\vec{p}}\right) \equiv \frac{1}{2m^2c^2}\frac{1}{r}\frac{d\hat{V}}{dr}\hat{\vec{S}}\cdot\left(\vec{r}\times\hat{\vec{p}}\right) = \frac{1}{2m^2c^2}\frac{1}{r}\frac{d\hat{V}}{dr}\hat{\vec{S}}\cdot\hat{\vec{L}}, \tag{12.484}$$

where $\hat{\vec{L}} = \vec{r}\times\hat{\vec{p}}$ is the orbital angular momentum.

The fifth term, $\hat{H}_{Darwin}$, arises from the fact that the electron is not a point particle, but spreads over a "volume" of the size of the Compton wavelength $\lambda \sim \hbar/mc$. In the case of an electron in an atom with Z protons, the Coulomb potential V is given by $V = -Ze^2/(4\pi\varepsilon_0 r)$ and using the fact that $\nabla^2(1/r) = -4\pi\delta(\vec{r})$ (see Appendix A), we can write

$$\nabla^2 V = -\frac{Ze^2}{4\pi\varepsilon_0}\nabla^2\left(\frac{1}{r}\right) = 4\pi\left(\frac{Ze^2}{4\pi\varepsilon_0}\right)\delta(\vec{r}) \tag{12.485}$$

and, hence, the Darwin term becomes

$$\hat{H}_{Darwin} = \frac{\hbar^2}{8m^2c^2}(\nabla^2\hat{V}) = \frac{Ze^2\hbar^2}{8m^2c^2\varepsilon_0}\delta(\vec{r}). \tag{12.486}$$

Due to $\delta(\vec{r})$, the Darwin term is non-zero only for the $S-$states. The Darwin term results from the quantum fluctuations in the position of the electron due to it high-frequency or rapid vibrations; this is known as the *Zitterbewegung*[70] effect. The frequency f of these vibrations is of the order of[71]

$$f \simeq \frac{mc^2}{\hbar} \simeq 10^{21}\ \mathrm{s}^{-1}. \tag{12.487}$$

It is noteworthy to mention that, while nonrelativistic quantum mechanics treats the above mentioned terms – the relativistic correction, the spin-orbit coupling, and the Darwin terms – in an ad hoc manner, these terms come out naturally from the Dirac equation and are treated without any need for perturbation theory.

As a reminder, the contributions of these terms to the energy were calculated in Chapter 9 (Subsection 9.2.3) using a first-order perturbation calculations the mean values of these terms for a hydrogen atom, as derived in Chapter 9, are given by

$$\langle nljm_j \mid \hat{H}_R \mid nljm_j\rangle = \langle -\frac{\hat{\vec{p}}^4}{8m^3c^2}\rangle = -\frac{\alpha^4 m_e c^2}{8n^4}\left(\frac{8n}{2l+1}-3\right) = -\frac{\alpha^2|E_n^{(0)}|}{4n^2}\left(\frac{8n}{2l+1}-3\right), \tag{12.488}$$

$$\langle nljm_j \mid \hat{H}_{SO} \mid nljm_j\rangle = \langle \frac{1}{2m^2c^2}\frac{1}{r}\frac{d\hat{V}}{dr}\hat{\vec{S}}\cdot\hat{\vec{L}}\rangle = \frac{\alpha^4 m_e c^2}{2n^3}\left[\frac{j(j+1)-l(l+1)-\frac{3}{4}}{l(l+1)(2l+1)}\right], \tag{12.489}$$

$$\langle n00 \mid \hat{H}_{Darwin} \mid n00\rangle = \langle \frac{e^2\hbar^2}{8m^2c^2\varepsilon_0}\delta(\vec{r})\rangle = \frac{1}{2n^3}m_e^2c^2\alpha^4, \tag{12.490}$$

where $j = l \pm 1/2$, $\alpha = 1/137$ is the fine structure constant and $E_n^{(0)}$ is the energy of the n^{th} excited state of the hydrogen atom, $E_n^{(0)} = -\alpha^2 m_e c^2/(2n^2)$. As shown in Chapter 9, the ratios of

[70] "Zitterbewegung" is a German word that means oscillatory or jittery motion.

[71] Using the fact that the rest mass-energy of an electron is $mc^2 \simeq 0.511$ MeV and $\hbar c \simeq 197.3$ MeVfm, we write the frequency as: $f = c/\lambda \simeq mc^2/\hbar \equiv (mc^2)c/(\hbar c) = (0.511\,\mathrm{MeV}) \times (10^8\,\mathrm{m/s})/(197.3\,\mathrm{MeVfm}) \simeq 10^{21}\ \mathrm{s}^{-1}$.

these terms to the energy of the hydrogen atom are of the order of α^2:

$$\frac{\langle \hat{H}_{SO} \rangle}{|E_n^{(0)}|} \simeq \frac{\alpha^2}{n}, \qquad \left| \frac{\langle \hat{H}_R \rangle}{E_n^{(0)}} \right| \simeq \frac{\alpha^2}{n}, \qquad \frac{\langle \hat{H}_D \rangle}{|E_n^{(0)}|} = \frac{\alpha^2}{n}. \tag{12.491}$$

For $n = 1$, all these terms are of the order of 10^{-4} since $\alpha^2 = (1/137)^2 \simeq 10^{-4}$.

Remarks
It is instructive to offer several contrasting remarks, at least from a pedagogical perspective, about the equations of Schrödinger, Pauli, Klein–Gordon, and Dirac.

- It can be argued that the Pauli equation lies somewhere between the Schrödinger and the Dirac equations, for it possesses and lacks features from both of them.

- The Pauli and Schrödinger equations are both nonrelativistic, whereas the Klein–Gordon and Dirac equations are relativistic.

- Both of the Pauli and Dirac equations account for spin, but the Schrödinger and Klein–Gordon equations do not deal, at least explicitly, with spin. However, and in contrast to the Dirac equation where the spin comes out naturally from the formalism, the Pauli equation introduces it in an *ad hoc* manner by simply replacing $\vec{p}^{\,2}$ with $(\vec{\sigma} \cdot \vec{p})^{\,2}$ in the Schrödinger equation. Although the Schrödinger equation does not deal with spin explicitly, it is not totally accurate to say that it is incapable of dealing with it. For instance, in the case of an electron, the Schrödinger equation does offer an accurate description of only one specific spin state at a time because the Shrödinger wave function has only one component (a scalar); as long as there is no external phenomenon (such as an external electromagnetic field) that can alter the orientation of the spin, the Schrödinger equation is equipped with the necessary theoretical tools to deal with that specific spin state. So, the Schrödinger equation is capable of offering an accurate description of a single spin state at a time.

- Unlike the Schrödinger and Klein–Gordon equations which deal with one component (scalar) wave functions and single component operators, the Pauli equation deals with two component spinors and 2×2 operators, while the Dirac equation has four component spinors and 4×4 operators. Since the Pauli theory deals with nonrelativistic spin 1/2 particles (no antiparticles), two component spinors are sufficient to describe the spin-up and spin-down states of the particle; the Dirac spinor describes jointly a particle and its antiparticle each requiring two components to account for the spin-up and spin-down states.

- While the Schrödinger and Pauli equations do not account for antiparticles, the Klein–Gordon and Dirac equations handle both particles and antiparticles.

- The Dirac and Pauli equations give the correct gyromagnetic ratio for the electron, whereas the Schödinger equation offers no prescription for it because it is a spinless equation.

Table 12.3: Some space-time symmetries and their corresponding conserved quantities.

Symmetry	Conserved quantity
Invariance under space translation	Linear momentum
Invariance under space rotations	Angular momentum
Invariance under time translation	Total energy

12.4.21 Discrete Symmetries of the Dirac Equation

Transformations, involving either spacetime or other degrees of freedom, play an important role in understanding the underlying symmetries of a theory; a symmetry is an operation which, when performed on a system, will leave its physical laws unchanged. Symmetries and conservation laws are interconnected. As shown in Table 12.3, which illustrates the conservation laws of some quantities such as liner momentum, energy, and angular momentum, and as Noether's theorem[72] states, *every symmetry has a corresponding conserved quantity*; interestingly, conservation laws can be verified experimentally. Symmetries can be best understood by studying how key observable quantities transform from one system of reference to another. *Transformations can be continuous or discrete*; we have seen a number of such symmetries in Chapters 2, 3, and 7. While continuous symmetries deal with continuous changes of a set of key parameters of the theory (e.g., infinitesimal rotations or translations), discrete symmetries deal with some kind of single, isolated reflections or inversions.

In the following three subsections, we are going to deal with the three most important discrete symmetries of the Dirac equation and their impact on its solutions: *charge conjugation (C), parity ($\mathcal{P}$), and time reversal (T)*. In short, charge conjugation deals with the interchange of particles and antiparticles, parity with spatial reflections, and time reversal with the reversal of time or, more precisely, the reversal of the direction of motion. After treating these three symmetries, we conclude this section with a discussion of the $C\mathcal{P}T$ theorem which states that, although each one of C, $\mathcal{P}$ and T symmetries may not be good symmetries individually (i.e., not conserved), their product $C\mathcal{P}T$ is always a good symmetry. In fact, the invariance of the Dirac equation under the $C\mathcal{P}T$ symmetry provides a justification for the Feynman-Stückelberg interpretation of negative energies.

12.4.22 Charge Conjugation and Covariance of the Dirac Equation

12.4.22.1 Charge Conjugation and Covariance of the Dirac Equation

In addition to giving us particle solutions, the Dirac equation predicts the existence of antiparticles (antimatter). Charge conjugation or charge inversion consists in replacing all particles with their partner antiparticles (and vice versa) by simply reversing the electric charges. In this subsection, we want to look at the charge conjugation symmetry and the invariance of the Dirac equation under it. We represent this symmetry by the charge conjugation operator $\hat{C}$ which we need to find.

Classically, charge conjugation can be represented by reversal charge: $q \xrightarrow{C} -q$. Additionally, since electric charge is the source of electric and magnetic fields, $\vec{E}$ and $\vec{B}$, we would

[72] Emmy Noether, "Invariante Variationsprobleme," (1918); translated from German by Mort Tavel, "Invariant Variation Problems," *Transport Theory and Statistical Physics*, **1** (3), 186–207 (1971).

need to reverse them: $\vec{E} \xrightarrow{C} -\vec{E}$ and $\vec{B} \xrightarrow{C} -\vec{B}$. However, the Maxwell equations are invariant under charge conjugation.

Quantum mechanically, charge conjugation, also called particle-antiparticle conjugation, affects only the *internal degrees of freedom*, contrary to the symmetries of parity and time reversal (to be treated below) which affect degrees of freedom that are external to the particle (the space-time coordinates). In addition to affecting electric charge, charge conjugation involves reversing all the internal quantum numbers such as lepton number, baryon number, and strangeness (we don't deal with lepton and baryon numbers nor with strangeness in this book); hence, a quantum state $\psi_{|Q|}(t, \vec{r})$, where $|Q|$ represents all the internal quantum numbers (electric charge q, lepton number, baryon number, etc: $|Q| = |q, \cdots|$), transforms like: $\psi_{|Q|}(t, \vec{r}) \xrightarrow{C} \psi_{|-Q|}(t, \vec{r})$. However, charge conjugation does not affect quantities such as mass, energy, momentum, spin, space coordinates, and time.

Covariance of the Dirac equation under charge conjugation

Consider an electron that is subjected to an electromagnetic field $A_\mu = (U/c, \vec{A})$, where U is the electric scalar potential and $\vec{A}$ is the magnetic vector potential. According to the *minimal coupling* principle, the Dirac equation for an *electron* interacting with an electromagnetic field A_μ can be obtained by making the following substitution (which can be obtained from (12.125) with $q = -e$)

$$p^\mu \longrightarrow p^\mu + eA^\mu \qquad \text{or} \qquad i\hbar\partial^\mu \longrightarrow i\hbar\partial^\mu + eA^\mu, \tag{12.492}$$

into the Dirac equation (12.231) for a *free* electron; this substitution yields the following equation:

$$\gamma_\mu \left(i\hbar\partial^\mu + eA^\mu\right) \psi(x^\mu) - mc\psi(x^\mu) = 0. \tag{12.493}$$

For a *positron* in the same electromagnetic field, the Dirac equation can be obtained from (12.493) by simply reversing the sign of the charge of the electron (i.e, $-e \longrightarrow +e$) and replacing $\psi(x^\mu)$ with its transform[73] $\psi_C(x^\mu)$:

$$\gamma_\mu \left(i\hbar\partial^\mu - eA^\mu\right) \psi_C(x^\mu) - mc\psi_C(x^\mu) = 0. \tag{12.494}$$

The problem now reduces to finding the charge conjugation operator $\hat{C}$ that will transform (12.493) into (12.494), thus making the Dirac equation covariant, since both of them have the same form. For this, we will need to start from (12.493) and work out the formalism until we end up with (12.494); this process will allow us to determine $\hat{C}$.

First, let us take the complex conjugate of (12.493)

$$\gamma_\mu^* \left(-i\hbar\partial^\mu + eA^\mu\right) \psi^*(x^\mu) - mc\psi^*(x^\mu) = 0. \tag{12.495}$$

Next, we assume that an operator $\hat{C}$ exists and that its matrix is non-singular; hence $\hat{C}^{-1}\hat{C} = I_4$. Now, replacing $\psi^*(x^\mu)$ with $(\hat{C}^{-1}\hat{C})\psi^*(x^\mu)$ in (12.495) and multiplying on the left by $\hat{C}$, we obtain

$$\hat{C}\gamma_\mu^* \left(-i\hbar\partial^\mu + eA^\mu\right) (\hat{C}^{-1}\hat{C})\psi^*(x^\mu) - mc\hat{C}(\hat{C}^{-1}\hat{C})\psi^*(x^\mu) = 0. \tag{12.496}$$

Now, using the fact that $\left(-i\hbar\partial^\mu + eA^\mu\right)\hat{C}^{-1} = -\hat{C}^{-1}\left(i\hbar\partial^\mu - eA^\mu\right)$ and $\hat{C}\hat{C}^{-1} = I_4$, we can reduce (12.496) to

$$-\hat{C}\gamma_\mu^*\hat{C}^{-1} \left(i\hbar\partial^\mu - eA^\mu\right) \hat{C}\psi^*(x^\mu) - mc\hat{C}\psi^*(x^\mu) = 0. \tag{12.497}$$

[73]Note that the charge conjugation operator $\hat{C}$ acts neither on time nor on the spatial coordinates, x^μ.

Finally, in order for (12.497) to be equivalent to (12.494), the charge conjugation operator $\hat{C}$ must satisfy the following two conditions

$$\psi_C(x^\mu) = \hat{C}\psi^*(x^\mu), \qquad \gamma_\mu = -\hat{C}\gamma_\mu{}^*\hat{C}^{-1}. \tag{12.498}$$

The second condition can be rewritten as

$$\gamma_\mu = -\hat{C}\gamma_\mu{}^*\hat{C}^{-1} \quad \Longrightarrow \quad \gamma_\mu\hat{C} = -\hat{C}\gamma_\mu{}^*\hat{C}^{-1}\hat{C} \quad \Longrightarrow \quad \gamma_\mu\hat{C} = -\hat{C}\gamma_\mu{}^*. \tag{12.499}$$

Now, since the γ^0, γ^1, and γ^3 matrices are real (i.e., $\gamma^{0^*} = \gamma^0$, $\gamma^{1^*} = \gamma^1$, $\gamma^{3^*} = \gamma^3$) while γ^2 is a purely imaginary matrix (i.e., $\gamma^{2^*} = -\gamma^2$), the only matrix that can satisfy $\gamma_\mu\hat{C} = -\hat{C}\gamma_\mu{}^*$ is γ^2; thus, we can reduce (12.499) to

$$\gamma_\mu\hat{C} = -\hat{C}\gamma_\mu{}^* \quad \Longrightarrow \quad \gamma^2\hat{C} = \hat{C}\gamma^2. \tag{12.500}$$

This yields $\hat{C} = \eta\gamma^2$ where η is an undetermined phase factor. However, since we know that $\hat{C}$ converts particles into antiparticles as will be illustrated in the calculation carried out below (see (12.508)), we can set $\eta = i$. Additionally, and invoking the matrix representation (12.223) of γ^2 and by taking $\eta = i$, we can easily verify that

$$\hat{C} = i\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{12.501}$$

satisfies (12.500). Since charge conjugation involves also taking the complex conjugate of the wave function, we would need to include a complex conjugation operator $\hat{K}$ in $\hat{C}$; hence we have

$$\boxed{\hat{C} = i\gamma^2\hat{K} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\hat{K}, \qquad \psi_C(x^\mu) = \hat{C}\psi(x^\mu) = i\gamma^2\psi^*(x^\mu).} \tag{12.502}$$

In summary, we have shown that the charge conjugation operator $\hat{C} = i\gamma^2\hat{K}$ transforms the Dirac equation for an electron into the Dirac equation for a positron; i.e., it transforms (12.493) into (12.494). We see that the spinors $\psi(x^\mu)$ and $\psi_C(x^\mu)$ describe particles that have the same mass but with opposite charges; so, *charge conjugation transforms a particle spinor $\psi(x^\mu)$ into its antiparticle spinor $\psi_C(x^\mu)$ and vice versa.*

Properties of the charge conjugation operator

- It is easy to verify that the matrix (12.501) is equal to its inverse: $\hat{C} = \hat{C}^{-1}$. This is expected, since applying $\hat{C}$ two successive times, we would end up with the original state. We can also obtain directly by calculating the inverse of (12.501):

$$\hat{C}^{-1} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{12.503}$$

- Applying $\hat{C}$ twice on the wave function, we end up with the original state

$$\hat{C}^2\psi(x^\mu) = \eta_C\hat{C}\psi_C(x^\mu) = \eta_C^2\psi(x^\mu) = \psi(x^\mu) \qquad \Longrightarrow \qquad \hat{C}^2 = I_4, \quad \eta_C = \pm 1, \tag{12.504}$$

where η_C is a phase factor known as the charge parity or simply C−parity. Similarly, using (12.501), we obtain the same result:

$$\hat{C}^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = I_4 \tag{12.505}$$

- Normalizing both $\psi(x^\mu)$ and $\psi_C(x^\mu)$, we obtain

$$1 = \langle\psi|\psi\rangle = \langle_C\psi|\psi_C\rangle = \langle\psi \mid \hat{C}^\dagger\hat{C} \mid \psi\rangle \qquad \Longrightarrow \qquad \hat{C}^\dagger\hat{C} = I_4. \tag{12.506}$$

- From (12.501), we can verify that $\hat{C}$ is Hermitian: $\hat{C}^\dagger = \hat{C}$. We can thus summaries the properties of $\hat{C}$ in the following relation

$$\hat{C}^2 = \hat{C} = \hat{C}^\dagger = \hat{C}^{-1}, \qquad \text{and} \qquad \hat{C}^\dagger\hat{C} = I_4. \tag{12.507}$$

Charge conjugation transforms particle spinors to their antiparticle spinors
Let us see how charge conjugation converts particle into an antiparticle states and vice versa. For this, we need to apply the operator (12.502) on the particle states (12.337) for a free particle.

So, applying $\hat{C}$ on $\psi_1(x^\mu)$ (i.e., applying (12.502) on (12.337)), we obtain

$$\begin{aligned} \hat{C}\psi_1(x^\mu) &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\hat{K}\psi_1(x^\mu) = N\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \end{pmatrix}^* e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar} \\ &= N\begin{pmatrix} \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix} e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar} \equiv v_1(p^\mu)e^{i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar} \\ &= \Phi_3(x^\mu), \end{aligned} \tag{12.508}$$

where $\Phi_3(x^\mu)$ is listed in (12.338); in deriving the above expression, we have used the fact that the normalization constant is real, $N = \sqrt{E^{(+)} + mc^2}$. Hence, we can write

$$\hat{C} \mid u_1\rangle^\uparrow = \mid v_1\rangle^\uparrow \qquad \longrightarrow \qquad \hat{C} \mid e^-\rangle^\uparrow = \mid e^+\rangle^\uparrow, \tag{12.509}$$

where the up-arrow (↑) stands for an upward spin, and (↓) for a downward spin. So, *under the charge conjugation operator, the particle spinor* $\mid u_1\rangle$ *transforms to the antiparticles spinor* $\mid v_1\rangle$; notice

that $|\, u_1\rangle$ has spin-up, charge $-e$, and positive energy $E^{(+)}$, while $|\, v_1\rangle$ has spin-up, charge $+e$, and positive energy.

Similarly, applying $\hat{C}$ on $\psi_2(x^\mu)$ of (12.337), we obtain

$$\begin{aligned}
\hat{C}\psi_2(x^\mu) &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \hat{K}\psi_1(x^\mu) = N \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ \frac{c(p_x - ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix}^* e^{i(E^{(+)}t - \vec{p}\cdot\vec{r})/\hbar} \\
&= N \begin{pmatrix} \frac{-cp_z}{E^{(+)}+mc^2} \\ -\frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \\ -1 \\ 0 \end{pmatrix} e^{i(E^{(+)}t - \vec{p}\cdot\vec{r})/\hbar} \equiv -v_2(p^\mu) e^{i(E^{(+)}t - \vec{p}\cdot\vec{r})/\hbar} \\
&= -\Phi_4(x^\mu), \qquad (12.510)
\end{aligned}$$

where $\Phi_4(x^\mu)$ is listed in (12.338). Hence, we can write

$$\hat{C}\,|\, u_2\rangle^\downarrow = -\,|\, v_2\rangle^\downarrow \qquad \longrightarrow \qquad \hat{C}\,|\, e^-\rangle^\downarrow = -\,|\, e^+\rangle^\downarrow. \qquad (12.511)$$

That is, the charge conjugation operator transforms the spinor, $|\, u_2\rangle$, of a positive energy, spin-down electron into the spinor, $|\, v_2\rangle$, of a positive energy, spin-down positron.

In summary, charge conjugation converts each particle into its antiparticle and vice versa:

$$\hat{C}\,|\, \psi\rangle = |\, \psi_A\rangle, \qquad (12.512)$$

where $|\, \psi\rangle$ describes a particle and $|\, \psi_A\rangle$ its antiparticle. Since $\hat{C}^2 = I_4$, the eigenvalues of $\hat{C}$ are $C = \pm 1$. However, most particle states are not eigenstates of charge conjugation; $\hat{C}$ changes, for instance, protons into antiprotons and positively charged pions into negative pions:

$$\hat{C}\;|\, p\rangle = |\, \overline{p}\rangle, \qquad \hat{C}\;|\, \pi^+\rangle = |\, \pi^-\rangle. \qquad (12.513)$$

Only particles that are their own antiparticles are eigenstates of $\hat{C}$:

$$\hat{C}\,|\, \psi\rangle = |\, \psi_A\rangle = \pm\,|\, \psi\rangle. \qquad (12.514)$$

The photon is such an example. If all the particle sources of an electromagnetic field are changed into their antiparticles, the electric charges change sign and so does the vector potential $\vec{A}$; hence, charge conjugation of a photon state $|\, \gamma\rangle$ changes sign as well

$$\hat{C}\,|\, \gamma\rangle = -\,|\, \gamma\rangle. \qquad (12.515)$$

As a result, the state of N photons, $|\, N\gamma\rangle$, transforms are follows

$$\hat{C}\,|\, N\gamma\rangle = (-1)^N\,|\, \gamma\rangle. \qquad (12.516)$$

For example, since a π^0 particle is known to decay electromagnetically into two photons (i.e., $\pi^0 \longrightarrow 2\gamma$), the charge conjugation of the pion π^0 is positive:

$$\hat{C}\,|\, \pi^0\rangle = |\, \pi^0\rangle. \qquad (12.517)$$

Finally, we should note that all interactions were thought to be invariant under charge conjugation. However, soon after the discovery of parity violation in $\beta-$decay, it was found that charge conjugation was also violated in weak interactions.

12.4.23 Parity and Covariance of the Dirac Equation

In this subsection, we want to look at the transformation of the Dirac equation under parity and at the intrinsic parity of the Dirac spinors. As seen in Chapter 2, the parity operator $\hat{\mathcal{P}}$ is defined as the spatial inversion of various quantities (vectors, wave functions, etcetera) through the origin of the reference system. In particular, we want to find out if the free particle Dirac equations is covariant under Parity transformations.

12.4.23.1 Parity Transformations in Classical and Quantum Physics

At this level, let us discursively take a short detour to examine how parity transforms several various quantities in classical physics and quantum mechanics.

Parity in Classical Physics
Classically, parity operation $\mathcal{P}$ is the reflection of the position vector $\vec{r}$ through the origin while leaving time intact:

$$\vec{r}\,' = \mathcal{P}\,\vec{r} = -\vec{r} \quad \text{or} \quad \vec{r} \xrightarrow{\mathcal{P}} -\vec{r}, \qquad t \xrightarrow{\mathcal{P}} t'; \tag{12.518}$$

the position vector is said to be odd under parity. Similarly, since time is not affected by spatial reflections, the linear momentum is also odd under parity

$$\vec{p}\,' = \mathcal{P}\vec{p} = m\frac{d(\mathcal{P}\vec{r})}{dt} = -m\frac{d\vec{r}}{dt} = -\vec{p}. \tag{12.519}$$

Clearly, the application of the parity operator two successive times would leave the spatial coordinates unchanged: $\mathcal{P}^2\,\vec{r} = \mathcal{P}\,(-\vec{r}) = \vec{r}$.

It is worth reminding that there are two kinds of vectors, *polar* and *axial*, and it is important to see how they transform under parity. First, *polar* vectors: they are simply the normal vectors such as the position and momentum vectors, $\vec{r}$ and $\vec{p}$, which are defined by their components such as polar coordinates (magnitude and direction or angle) or their Cartesian coordinate. *Under parity, the components of polar vectors get transformed to their respective negative values* as seen in (12.518) and (12.519). Second, we have *axial* vectors, also called *pseudo-vectors*. Examples of axial vectors are cross products of polar vectors such as the orbital angular momentum $\vec{L} = \vec{r}\times\vec{p}$. In contrast to polar vectors, *axial vectors remain unchanged under spatial inversion*, since the cross product of two vectors is invariant (or even) under parity:

$$\vec{L} = \vec{r}\times\vec{p} \xrightarrow{\mathcal{P}} \vec{L}\,' = \mathcal{P}(\vec{L}) = (-\vec{r})\times(-\vec{p}) \equiv \vec{L}. \tag{12.520}$$

We should note that *scalars are invariant under parity*; for instance, the square of a vector, which is a scalar, remains unchanged under parity

$$\vec{r}^{\,2} \xrightarrow{\mathcal{P}} \vec{r}^{\,\prime 2} = \mathcal{P}\vec{r}^{\,2} = (-\vec{r})^{\,2} = \vec{r}^{\,2} \equiv r^2 \quad \text{or} \quad r'^{\,2} = r^2. \tag{12.521}$$

There is another class of scalars which are called *pseudo-scalars* such as the scalar product between a vector $\vec{A}$ and the cross product between two other vectors $\vec{B}\times\vec{C}$. Clearly, *pseudo-scalars transform oppositely from scalars*:

$$\vec{A}\cdot(\vec{B}\times\vec{C}) \xrightarrow{\mathcal{P}} (-\vec{A})\cdot\left[(-\vec{B})\times(-\vec{C})\right] = -\vec{A}\cdot(\vec{B}\times\vec{C}). \tag{12.522}$$

Finally, the transformation of 4-vectors under parity can be easy obtained by analogy to (12.518) and (12.519); the position and momentum 4-vectors transform as follows:

$$x'^{\mu} = \mathcal{P}x^{\mu} \equiv (x^0, -\vec{r}) \longrightarrow x'^0 = x^0 \quad \text{or} \quad t' = t \quad \text{and} \quad \vec{r}\,' = -\vec{r}, \tag{12.523}$$

$$p'^{\mu} = \mathcal{P}p^{\mu} \equiv (p^0, -\vec{p}) \longrightarrow p'^0 = p^0 \quad \text{or} \quad E' = E \quad \text{and} \quad \vec{p}\,' = -\vec{p}. \tag{12.524}$$

We can also express these spatial reflections by means of a Lorentz transformation $\Lambda^{\mu}{}_{\nu}$:

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu}, \qquad p'^{\mu} = \Lambda^{\mu}{}_{\nu}p^{\nu}, \qquad \text{where} \qquad \Lambda \equiv \left(\Lambda^{\mu}{}_{\nu}\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{12.525}$$

Parity in Quantum Mechanics

Quantum mechanically and by analogy to the classical expressions (12.518) and (12.519), the parity operator $\hat{\mathcal{P}}$ transforms the position operator $\hat{\vec{r}}$ and linear momentum operator $\hat{\vec{p}}$ as follows:

$$\hat{\vec{r}}\,' = \hat{\mathcal{P}}^{\dagger}\,\hat{\vec{r}}\,\hat{\mathcal{P}} = -\hat{\vec{r}} \quad \text{or} \quad \hat{\vec{r}} \xrightarrow{\hat{\mathcal{P}}} -\hat{\vec{r}}, \tag{12.526}$$

$$\hat{\vec{p}}\,' = \hat{\mathcal{P}}^{\dagger}\,\hat{\vec{p}}\,\hat{\mathcal{P}} = -\hat{\vec{p}} \quad \text{or} \quad \hat{\vec{p}} \xrightarrow{\hat{\mathcal{P}}} -\hat{\vec{p}}. \tag{12.527}$$

Similarly, and by analogy to (12.520), parity leaves the orbital angular momentum operator $\hat{\vec{L}}$ unchanged:

$$\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}} \xrightarrow{\hat{\mathcal{P}}} \hat{\vec{L}}\,' = \mathcal{P}^{\dagger}\,\hat{\vec{L}}\,\mathcal{P} = \left(\hat{\mathcal{P}}^{\dagger}\,\hat{\vec{r}}\,\hat{\mathcal{P}}\right) \times \left(\hat{\mathcal{P}}^{\dagger}\,\hat{\vec{p}}\,\hat{\mathcal{P}}\right) = (-\hat{\vec{r}}) \times (-\hat{\vec{p}}) \equiv \hat{\vec{L}}. \tag{12.528}$$

Note that the spin as well as the total angular momentum operators, $\hat{\vec{S}}$ and $\hat{\vec{J}}$, are also invariant under parity.

Worthy of note is the fact that *the basic commutation relations of quantum mechanics are invariant under parity* (i.e., under $\hat{\vec{r}} \xrightarrow{\hat{\mathcal{P}}} -\hat{\vec{r}}$, $\hat{\vec{p}} \xrightarrow{\hat{\mathcal{P}}} -\hat{\vec{p}}$, and $\hat{\vec{J}} \xrightarrow{\hat{\mathcal{P}}} \hat{\vec{J}}$):

$$[\hat{x}_j, \hat{p}_k] = i\hbar\delta_{jk} \xrightarrow{\hat{\mathcal{P}}} [\hat{x}'_j, \hat{p}'_k] = [-\hat{x}_j, -\hat{p}_k] \equiv [\hat{x}_j, \hat{p}_k] = i\hbar\delta_{jk}, \qquad j,k = x,y,z, \tag{12.529}$$

$$[\hat{J}_j, \hat{J}_k] = i\hbar\epsilon_{jkl}\hat{J}_l \xrightarrow{\hat{\mathcal{P}}} [\hat{J}'_j, \hat{J}'_k] = [\hat{J}_j, \hat{J}_k] = i\hbar\epsilon_{jkl}\hat{J}_l, \qquad j,k,l = x,y,z. \tag{12.530}$$

Hence, and as will be shown below, parity can be represented by a unitary operator.

As for wave functions, they transform under parity like:

$$\psi(x^{\mu}) \xrightarrow{\hat{\mathcal{P}}} \psi_P(x'^{\mu}) = \hat{\mathcal{P}}\psi(x^{\mu}) = \psi(x^0, -\vec{r}) = \psi(t, -\vec{r}), \tag{12.531}$$

where $\hat{\mathcal{P}}$ is a 4×4 matrix, since $\psi_P(x'^{\mu})$ and $\psi(x^{\mu})$ are 4-component spinors. Note that, the application of the parity operator two successive times would leave the spatial coordinates unchanged:

$$\hat{\mathcal{P}}^2\psi(x^{\mu}) = \hat{\mathcal{P}}\psi(x^0, -\vec{r}) = \psi(x^0, \vec{r}) \equiv \psi(x^{\mu}) \implies \hat{\mathcal{P}}^2 = I_4 \implies \hat{\mathcal{P}}^{-1} = \hat{\mathcal{P}}. \tag{12.532}$$

Hence, *the eigenvalues of the parity operator are*: $\mathcal{P} = \pm 1$. Additionally, parity preserves the norm of the wave function

$$\langle \psi_P | \psi_P \rangle = \langle \psi | \hat{\mathcal{P}}^\dagger \hat{\mathcal{P}} | \psi \rangle \equiv \langle \psi | \psi \rangle \quad \Longrightarrow \quad \hat{\mathcal{P}}^\dagger \hat{\mathcal{P}} = I_4. \tag{12.533}$$

Now, from (12.532) and (12.533), it is easy to verify that $\hat{\mathcal{P}}$ is hermitian

$$\hat{\mathcal{P}}^2 = I_4 \quad \text{and} \quad \hat{\mathcal{P}}^\dagger \hat{\mathcal{P}} = I_4 \quad \Longrightarrow \quad \hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}}. \tag{12.534}$$

Similarly, using (12.532) and (12.534), we can easily see that $\hat{\mathcal{P}}$ is a unitary operator

$$\hat{\mathcal{P}}^{-1} = \hat{\mathcal{P}} \quad \text{and} \quad \hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}} \quad \Longrightarrow \quad \hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}}^{-1} = \hat{\mathcal{P}} \quad \Longrightarrow \quad \hat{\mathcal{P}} \text{ is unitary.} \tag{12.535}$$

Note that we can write any state $|\psi\rangle$ as

$$|\psi\rangle = \frac{1}{2}(\hat{\mathcal{P}} + 1)\,|\psi\rangle - \frac{1}{2}(\hat{\mathcal{P}} - 1)|\psi\rangle \equiv |\psi\rangle_+ + |\psi\rangle_-, \tag{12.536}$$

where

$$|\psi\rangle_+ = \frac{1}{2}(\hat{\mathcal{P}} + 1)\,|\psi\rangle, \quad |\psi\rangle_- = -\frac{1}{2}(\hat{\mathcal{P}} - 1)\,|\psi\rangle, \tag{12.537}$$

Since $\hat{\mathcal{P}}^2 = I_4$ and using (12.537), we can easily show that: $\hat{\mathcal{P}}\,|\psi\rangle_\pm = \pm\,|\psi\rangle_\pm$:

$$\hat{\mathcal{P}}\,|\psi\rangle_+ = \frac{1}{2}(\hat{\mathcal{P}}^2 + \hat{\mathcal{P}})\,|\psi\rangle = \frac{1}{2}(I_4 + \hat{\mathcal{P}})\,|\psi\rangle = \frac{1}{2}(\hat{\mathcal{P}} + 1)\,|\psi\rangle = |\psi\rangle_+, \tag{12.538}$$

$$\hat{\mathcal{P}}\,|\psi\rangle_- = -\frac{1}{2}(\hat{\mathcal{P}}^2 - \hat{\mathcal{P}})\,|\psi\rangle \equiv -\frac{1}{2}(I_4 - \hat{\mathcal{P}})\,|\psi\rangle = \frac{1}{2}(\hat{\mathcal{P}} - 1)\,|\psi\rangle = -\,|\psi\rangle_-. \tag{12.539}$$

Hence, the states $|\psi\rangle_+$ and $|\psi\rangle_-$ are eigenstates of $\hat{\mathcal{P}}$ with eigenvalues ± 1, respectively.

12.4.23.2 Covariance of the Dirac Equation Under Parity

Let us now look at the effect of parity inversion on the Dirac equation; for this, let us rewrite (12.230) as follows

$$\frac{i\hbar}{c}\gamma^0 \frac{\partial \psi(t, \vec{r})}{\partial t} = \left(-i\hbar\gamma^1 \frac{\partial}{\partial x} - i\hbar\gamma^2 \frac{\partial}{\partial y} - i\hbar\gamma^3 \frac{\partial}{\partial z} + mc\right)\psi(t, \vec{r}). \tag{12.540}$$

The problem now reduces to finding the form of the parity operator $\hat{\mathcal{P}}$ that will make the Dirac equation (12.540) covariant under spatial reflection. That is, we need to find an operator $\hat{\mathcal{P}}$ that will make the transformed state $\psi_P(t', \vec{r}') = \hat{\mathcal{P}}\psi(t, \vec{r})$ a solution of the following equation:

$$\frac{i\hbar}{c}\gamma^0 \frac{\partial \psi_P(t', \vec{r}')}{\partial t'} = \left(-i\hbar\gamma^1 \frac{\partial}{\partial x'} - i\hbar\gamma^2 \frac{\partial}{\partial y'} - i\hbar\gamma^3 \frac{\partial}{\partial z'} + mc\right)\psi_P(t', \vec{r}'); \tag{12.541}$$

this equation is the parity transform of (12.540); notice that (12.541) and (12.540) have the same form. In our approach, we are going to start with (12.540) and work our way until we end up with (12.541) and in the process find the form of $\hat{\mathcal{P}}$.

First, let us begin by inserting the state $\psi(t, \vec{r}) = \hat{\mathcal{P}}^{-1}\psi_P(t', \vec{r}')$ into (12.540):

$$\frac{i\hbar}{c}\gamma'^0\hat{\mathcal{P}}^{-1}\frac{\partial\psi_P(t', \vec{r}')}{\partial t} = \left(-i\hbar\gamma'^1\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial x} - i\hbar\gamma'^2\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial y} - i\hbar\gamma'^3\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial z} + mc\,\hat{\mathcal{P}}^{-1}\right)\psi_P(t', \vec{r}'). \quad (12.542)$$

Next, since $t = t'$ and $\vec{r} = -\vec{r}'$, we have

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t'} \qquad \frac{\partial}{\partial x} = -\frac{\partial}{\partial x'} \qquad \frac{\partial}{\partial y} = -\frac{\partial}{\partial y'} \qquad \frac{\partial}{\partial z} = -\frac{\partial}{\partial z'}; \quad (12.543)$$

inserting these expressions into (12.542), we obtain

$$\frac{i\hbar}{c}\gamma'^0\hat{\mathcal{P}}^{-1}\frac{\partial\psi_P(t', \vec{r}')}{\partial t'} = \left(i\hbar\gamma'^1\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial x'} + i\hbar\gamma'^2\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial y'} + i\hbar\gamma'^3\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial z'} + mc\,\hat{\mathcal{P}}^{-1}\right)\psi_P(t', \vec{r}'). \quad (12.544)$$

Now, a multiplication of all terms of this equation on the left with γ'^0 leads to

$$\frac{i\hbar}{c}\gamma'^0\gamma'^0\hat{\mathcal{P}}^{-1}\frac{\partial\psi_P(t', \vec{r}')}{\partial t'} = \left(i\hbar\gamma'^0\gamma'^1\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial x'} + i\hbar\gamma'^0\gamma'^2\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial y'} + i\hbar\gamma'^0\gamma'^3\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial z'} + mc(\gamma'^0\hat{\mathcal{P}}^{-1})\right)\psi_P(t', \vec{r}'). \quad (12.545)$$

Since $\gamma'^0\gamma'^j = -\gamma'^j\gamma'^0$, where $j = 1, 2, 3$, we can reduce this relation to

$$\frac{i\hbar}{c}\gamma'^0\gamma'^0\hat{\mathcal{P}}^{-1}\frac{\partial\psi_P(t', \vec{r}')}{\partial t'} = \left(-i\hbar\gamma'^1\gamma'^0\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial x'} - i\hbar\gamma'^2\gamma'^0\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial y'} - i\hbar\gamma'^3\gamma'^0\hat{\mathcal{P}}^{-1}\frac{\partial}{\partial z'} + mc\gamma'^0\hat{\mathcal{P}}^{-1}\right)\psi_P(t', \vec{r}'). \quad (12.546)$$

Clearly, if $\gamma'^0\hat{\mathcal{P}}^{-1} = \pm I_4$, we can reduce this equation to (12.541). Hence, *the Dirac equation is covariant under parity* provided we have $\gamma'^0\hat{\mathcal{P}}^{-1} = \pm I_4$.

Now, using the fact that γ'^0 is equal to its inverse, the condition $\gamma'^0\hat{\mathcal{P}}^{-1} = \pm I_4$ yields the parity operator $\hat{\mathcal{P}} = \pm\gamma'^0$ which can be written in matrix form as follows:

$$\hat{\mathcal{P}} = \pm\gamma'^0 = \pm\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (12.547)$$

So, the Dirac equation is covariant under parity provided that the parity operator is given by $\hat{\mathcal{P}} = \pm\gamma'^0$ and, hence, the Dirac spinors would transform as follows

$$\boxed{|\psi\rangle \longrightarrow \hat{\mathcal{P}}|\psi\rangle = \pm\gamma'^0|\psi\rangle.} \quad (12.548)$$

The eigenvalues $P = \pm 1$ of the parity operator $\hat{\mathcal{P}}$ are phase factors; hence, the transformed wave function $|\psi_P\rangle$ differs from the original wave function $|\psi\rangle$ by only a phase factor: $|\psi_P\rangle = \pm\gamma'^0|\psi\rangle$. The quantum number $P = \pm 1$ is called the *intrinsic parity* of the particle; notice that the intrinsic parity can be defined if a particle is at rest. If we have a system system of particles, its intrinsic parity is given by the product of the intrinsic parities of the particles.

We should mention that, using $\hat{\mathcal{P}} = \pm\gamma'^0$, we have shown in Problem 12.26 on Page 858 that the *intrinsic parity of a particle at rest is opposite to that of an antipartilcle at rest*:

$$\hat{\mathcal{P}}|u_j\rangle = \pm|u_j\rangle, \qquad \hat{\mathcal{P}}|v_j\rangle = \mp|v_j\rangle, \qquad j = 1, 2, \quad (12.549)$$

where $|u_1\rangle$ and $|u_2\rangle$ are particle spinors, while $|v_1\rangle$ and $|v_2\rangle$ are antiparticle spinors. This is a consequence of the Dirac equation in that the intrinsic parities of a fermion (F) and its antifermion (AF) obey this relation

$$\hat{\mathcal{P}}_F \hat{\mathcal{P}}_{AF} = -I_4. \tag{12.550}$$

By convention, fermions are chosen to have positive parity and antifermions negative parity:

$$\hat{\mathcal{P}}_F = +\gamma^0, \qquad \hat{\mathcal{P}}_{AF} = -\gamma^0. \tag{12.551}$$

According to this convention, equation (12.549) becomes

$$\hat{\mathcal{P}}|u_j\rangle = +|u_j\rangle, \qquad \hat{\mathcal{P}}|v_j\rangle = -|v_j\rangle, \qquad j = 1, 2. \tag{12.552}$$

Example 12.10

Using the algebraic properties of the γ−matrices, verify that the 4-current $j^\mu = \overline{\psi}\gamma^\mu\psi$ transforms as a vector under parity; i.e., show that $j^0 \longrightarrow j^0$ and $\vec{j} \longrightarrow -\vec{j}$.

Solution

As shown in Eq. (12.281), since $\gamma^0\gamma^0 = I_4$ and since $j^0 = c\overline{\psi}\gamma^0\psi = c\psi^\dagger\gamma^0\gamma^0\psi = c\psi^\dagger I_4\psi$, we can write the transformation of j^0 under parity as follows

$$j^0 \xrightarrow{\hat{\mathcal{P}}} \hat{\mathcal{P}}^\dagger j^0 \hat{\mathcal{P}} = c\,\hat{\mathcal{P}}^\dagger \left(\psi^\dagger I_4\psi\right)\hat{\mathcal{P}} = c\,\psi^\dagger \left(\hat{\mathcal{P}}^\dagger I_4\hat{\mathcal{P}}\right)\psi = c\psi^\dagger\psi = j^0, \tag{12.553}$$

where we have used the fact that the parity operator is unitary, $\hat{\mathcal{P}}^\dagger I_4\hat{\mathcal{P}} = \hat{\mathcal{P}}^\dagger\hat{\mathcal{P}} = I_4$.

Next, from Eq. (12.281) and since, as displayed in (12.547), the parity operator is given by $\hat{\mathcal{P}} = \pm\gamma^0$ and $\vec{j} = c\psi^\dagger\gamma^0\vec{\gamma}\psi$, we can write

$$\vec{j} \xrightarrow{\hat{\mathcal{P}}} \hat{\mathcal{P}}^\dagger\vec{j}\hat{\mathcal{P}} = c\psi^\dagger \left(\hat{\mathcal{P}}^\dagger\gamma^0\vec{\gamma}\hat{\mathcal{P}}\right)\psi = c\psi^\dagger \left((\gamma^0)^\dagger\gamma^0\vec{\gamma}\gamma^0)\right)\psi. \tag{12.554}$$

Now, since $(\gamma^0)^\dagger = \gamma^0$ and since $\vec{\gamma}\gamma^0 = -\gamma^0\vec{\gamma}$ (Eq. (12.180) yields $\gamma^j\gamma^0 = -\gamma^0\gamma^j$ where $j = 1, 2, 3$), we can reduce (12.554) to

$$\hat{\mathcal{P}}^\dagger\vec{j}\hat{\mathcal{P}} = c\psi^\dagger \left((\gamma^0)^\dagger\gamma^0\vec{\gamma}\gamma^0)\right)\psi = -c\psi^\dagger \left(\gamma^0\gamma^0\gamma^0\vec{\gamma})\right)\psi = -c\psi^\dagger \left(\gamma^0\vec{\gamma})\right)\psi = -\vec{j}, \tag{12.555}$$

where we have used the fact that $\gamma^0\gamma^0 = I_4$.

From (12.553) and (12.555), we see that while j^0 is invariant under parity, $\vec{j}$ transforms as a vector; namely:

$$j^0 \xrightarrow{\hat{\mathcal{P}}} j^0, \qquad \vec{j} \xrightarrow{\hat{\mathcal{P}}} -\vec{j}. \tag{12.556}$$

This was expected, since j^0 is a scalar and $\vec{j}$ is a vector. In Problem 12.31, we will show how to verify these two transformations using the matrices of γ^0, γ^1, and γ^2, γ^3, and $\hat{\mathcal{P}}$.

12.4.24 Time-Reversal and Covariance of the Dirac Equation

Time reversal deals essentially with the symmetry of the physical laws under the operation of reversing the direction of the time axis: $t \longrightarrow t' = -t$. Let us examine how time reversal applies to classical and quantum mechanical systems.

Time Reversal in Classical Physics

Classically, if $\vec{r}(t)$ is a solution to Newton's second law $\vec{F} = md^2\vec{r}/dt^2$, then also $\vec{r}(-t)$ is a solution. The operation of time inversion on some classical quantities works as follows

$$t \xrightarrow{T} t' = -t, \qquad \vec{r} \xrightarrow{T} \vec{r}\,' = \vec{r}, \tag{12.557}$$

$$\vec{p} \xrightarrow{T} \vec{p}\,' = m\frac{d\vec{v}}{dt'} = -m\frac{d\vec{v}}{dt} \equiv -\vec{p}, \tag{12.558}$$

$$\vec{L} = \vec{r}\times\vec{p} \xrightarrow{T} \vec{L}\,' = \vec{r}\,'\times\vec{p}\,' = \vec{r}\times(-\vec{p}) = -\vec{r}\times\vec{p} \equiv -\vec{L}. \tag{12.559}$$

So, time reversal changes the sign of quantities such as time, velocity, linear momentum and angular momentum (orbital and spin), but keeps the sign of spatial quantities (such as position vectors) and those quantities that depend on even powers of time (such as acceleration, force, energy, etc) intact.

Time Reversal in Quantum Mechanics

Under time reversal, 4-vectors transform like

$$x^\mu = (ct, \vec{r}) \xrightarrow{T} x'^\mu = (ct', \vec{r}\,') = (-ct, \vec{r}), \tag{12.560}$$

or, by means of a Lorentz transformation $\Lambda^\mu{}_\nu$, like

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \qquad \text{where} \qquad \Lambda \equiv \left(\Lambda^\mu{}_\nu\right) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{12.561}$$

By analogy to (12.557)–(12.559), we can infer the time reversal transformations of the position, momentum, and angular momentum operators:

$$\hat{\vec{r}} \xrightarrow{T} \hat{\vec{r}}\,' = \hat{T}^\dagger\hat{\vec{r}}\,\hat{T} \equiv \hat{\vec{r}}, \qquad \hat{\vec{p}} \xrightarrow{T} \hat{\vec{p}}\,' = \hat{T}^\dagger\hat{\vec{p}}\,\hat{T} \equiv -\hat{\vec{p}}, \qquad \hat{\vec{J}} \xrightarrow{T} \hat{\vec{J}}\,' = \hat{T}^\dagger\hat{\vec{J}}\,\hat{T} \equiv -\hat{\vec{J}}, \tag{12.562}$$

where $\hat{\vec{J}}$ is the most general form of an angular momentum operator (which can be a spin or orbital angular momentum or the sum of both).

The Time Reversal Operator cannot be Unitary

First, one can see that the time inversion relations (12.562) cannot be satisfied by a unitary operator, for if we use (12.562), *the basic commutation relations of quantum mechanics are not* invariant under the classical definition of time inversion:

$$[\hat{x}_j, \hat{p}_k] = i\hbar\delta_{jk} \xrightarrow{\hat{T}} \hat{T}^\dagger[\hat{x}_j, \hat{p}_k]\hat{T} = \hat{T}^\dagger(i\hbar\delta_{jk})\hat{T} \implies -[\hat{x}_j, \hat{p}_k] = i\hbar\delta_{jk}, \qquad j,k = x,y,z. \tag{12.563}$$

This can be shown as follows. On the one hand, since the quantity $i\hbar\delta_{jk}$ is a number and if $\hat{T}$ is unitary (i.e., $\hat{T}^\dagger\,\hat{T} = I_4$), we would be able to pull $\hat{T}$ to the left and this would lead to: $\hat{T}^\dagger(i\hbar\delta_{jk})\hat{T} = \hat{T}^\dagger\hat{T}(i\hbar\delta_{jk}) \equiv i\hbar\delta_{jk}$. On the other hand, using the relations $\hat{T}^\dagger\hat{x}_j\hat{T} = \hat{x}_j$ and $\hat{T}^\dagger\hat{p}_j\hat{T} = -\hat{p}_j$, we can write

$$\hat{T}^\dagger[\hat{x}_j,\,\hat{p}_k]\hat{T} = \hat{T}^\dagger\left(\hat{x}_j\hat{p}_k - \hat{p}_k\hat{x}_j\right)\hat{T} = (\hat{T}^\dagger\hat{x}_j\hat{T})(\hat{T}^\dagger\hat{p}_k\hat{T}) - (\hat{T}^\dagger\hat{p}_k\hat{T})(\hat{T}^\dagger\hat{x}_j\hat{T}) = -\hat{x}_j\,\hat{p}_k + \hat{p}_k\hat{x}_j = -[\hat{x}_j,\,\hat{p}_k]. \tag{12.564}$$

Similarly, using the relation $\hat{T}^\dagger\hat{J}_j\hat{T} = -\hat{J}_j$ which leads to $\hat{T}^\dagger(i\hbar\epsilon_{jkl}\hat{J}_l)\hat{T} = -i\hbar\epsilon_{jkl}\hat{J}_l$ and $\hat{T}^\dagger[\hat{J}_j,\,\hat{J}_k]\hat{T} = [-\hat{J}_j,\,-\hat{J}_k] = [\hat{J}_j,\,\hat{J}_k]$, it is easy to see that the angular momentum commutator is also not invariant

$$[\hat{J}_j,\,\hat{J}_k] = i\hbar\epsilon_{jkl}\hat{J}_l \xrightarrow{\hat{T}} \hat{T}^\dagger[\hat{J}_j,\,\hat{J}_k]\hat{T} = \hat{T}^\dagger(i\hbar\epsilon_{jkl}\hat{J}_l)\hat{T} \Longrightarrow [\hat{J}_j,\,\hat{J}_k] = -i\hbar\epsilon_{jkl}\hat{J}_l, \qquad j,k,l = x,y,z. \tag{12.565}$$

Hence, *time reversal cannot be represented by a unitary operator* as it does not preserve the fundamental commutators of quantum mechanics; i.e., $\hat{T}^\dagger \neq \hat{T}$ or $\hat{T}^\dagger\hat{T} \neq I_4$. This has to do with the above erroneous approach of having used the classical definition of time and tried to graft it on relativistic quantum mechanics. As a reminder, in non-relativistic quantum mechanics, as in classical physics, time is an *absolute parameter* which is independent of the spatial coordinates or reference frame. In relativistic quantum mechanics, however, time is not universal (not an absolute parameter), it is known as the *proper time*[74] which results from the reading of a clock in the rest-frame of the particle; hence, proper time changes differently along different time lines. So, in relativistic quantum mechanics, time reversal cannot be defined by the simple transformation $t \longrightarrow -t$ along with the use of a unitary transformation.

The way out of this problem is to use an *antiunitary* operator to represent $\hat{T}$. For this purpose, we have already covered the properties of antiunitary operators in Chapter 2 (see Subsection 2.4.8) where we have shown that *any antiunitary operator can be decomposed into the product of a unitary operator times a complex conjugation operator*. So, the way out is simply to write, as per *Wigner's theorem*[75], the time reversal operator $\hat{T}$ as $\hat{T} = \hat{U}\hat{K}$ where $\hat{U}$ is a unitary operator and $\hat{K}$ is an antiunitary complex conjugation operator that satisfies this condition: $\hat{K} = \hat{K}^\dagger = \hat{K}^{-1}$. In this way, the imaginary number "i" on the right-hand sides of the commutators in (12.563) and (12.565) would change sign to "$-i$" when we pull $\hat{T}$ to the left: i.e., $\hat{T}^\dagger i\hat{T} = -i\hat{T}^\dagger\hat{T}$. This would then make the commutators (12.563)–(12.565) invariant, since

$$\hat{T}^\dagger[\hat{x}_j,\,\hat{p}_k]\hat{T} = (\hat{T}^\dagger i\hat{T})\hbar\delta_{jk} \xrightarrow{\hat{T}} [\hat{x}'_j,\,\hat{p}'_k] = i\hbar\delta_{jk}, \qquad j,k = x,y,z, \tag{12.566}$$

and

$$\hat{T}^\dagger[\hat{J}_j,\,\hat{J}_k]\hat{T} = \hbar\epsilon_{jkl}(\hat{T}^\dagger i\hat{J}_l\hat{T}) \xrightarrow{\hat{T}} [\hat{J}'_j,\,\hat{J}'_k] = i\hbar\epsilon_{jkl}\hat{J}'_l, \qquad j,k,l = x,y,z. \tag{12.567}$$

Hence, in order to preserve the invariance of physical laws under time reversal, the operator $\hat{T}$ must include two operations: (a) *taking the complex conjugate* and (b) *inverting time* (i.e., let

[74] In the theory of special relativity, the proper time τ is the time in the particle's rest frame; it is related to an observer's time t by $\tau = t \times \sqrt{1 - v^2/c^2}$. Time τ is an invariant parameter as it is not specific to any coordinate system; it is a Lorentz scalar.

[75] According to Wigner theorem, a symmetry transformation S is represented either by a unitary operator, $\hat{S} = \hat{U}$, or an antiunitary operator, $\hat{S} = \hat{U}\hat{K}$, where $\hat{U}$ is unitary, and $\hat{K}$ denotes a complex conjugation operator.

$t \longrightarrow t' = -t$). As a result, the time inversion of the wave function $\psi(t, \vec{r})$ is not given by $\psi(-t, \vec{r})$, but by $\psi^*(-t, \vec{r})$:

$$\psi_T(t', \vec{r}') = \hat{T}\psi(t, \vec{r}) = \psi^*(-t, \vec{r}). \tag{12.568}$$

As we will see below, $\psi(t, \vec{r})$ is not a solution of the Dirac equation but $\psi^*(-t, \vec{r})$ is.

In summary, *contrary to space reflection, which is represented by a unitary operator, time inversion has to be represented by an antiunitary operator.* The task now reduces to finding the structure of the time reversal operator $\hat{T}$.

Finding the Time Reversal Operator that leaves the Dirac Equation Covariant

In what follows, we are going to determine the time reversal operator $\hat{T}$ that ensures the covariance of the Dirac equation (see (12.540)):

$$\frac{i\hbar}{c}\gamma^0\frac{\partial}{\partial t}\psi(x^\mu) = \left(-i\hbar\vec{\gamma}\cdot\vec{\nabla} + mc\right)\psi(x^\mu). \tag{12.569}$$

For this, let us take the complex conjugate of (12.569)

$$-\frac{i\hbar}{c}\gamma^{0^*}\frac{\partial\psi^*(t, \vec{r})}{\partial t} = \left(i\hbar\gamma^{1^*}\frac{\partial}{\partial x} + i\hbar\gamma^{2^*}\frac{\partial}{\partial y} + i\hbar\gamma^{3^*}\frac{\partial}{\partial z} + mc\right)\psi^*(t, \vec{r}). \tag{12.570}$$

From the definitions (12.222) – (12.223), we can easily obtain the complex conjugates of the γ^μ matrices (only γ^2 changes sign under complex conjugation):

$$\gamma^{0^*} = \gamma^0, \qquad \gamma^{1^*} = \gamma^1, \qquad \gamma^{2^*} = -\gamma^2, \qquad \gamma^{3^*} = \gamma^3, \tag{12.571}$$

which, when inserted into (12.570), lead to

$$-\frac{i\hbar}{c}\gamma^0\frac{\partial\psi^*(t, \vec{r})}{\partial t} = \left(i\hbar\gamma^1\frac{\partial}{\partial x} - i\hbar\gamma^2\frac{\partial}{\partial y} + i\hbar\gamma^3\frac{\partial}{\partial z} + mc\right)\psi^*(t, \vec{r}). \tag{12.572}$$

Multiplying both sides of this equation on the left by $\gamma^1\gamma^3$, we obtain

$$-\frac{i\hbar}{c}\left(\gamma^1\gamma^3\gamma^0\right)\frac{\partial\psi^*(t, \vec{r})}{\partial t} = \left[i\hbar\left(\gamma^1\gamma^3\gamma^1\right)\frac{\partial}{\partial x} - i\hbar\left(\gamma^1\gamma^3\gamma^2\right)\frac{\partial}{\partial y} + i\hbar\left(\gamma^1\gamma^3\gamma^3\right)\frac{\partial}{\partial z} + mc\gamma^1\gamma^3\right]\psi^*(t, \vec{r}). \tag{12.573}$$

Now, using the anticommutator $\left\{\gamma^i, \gamma^j\right\} = -2\delta_{ij}$ along with $(\gamma^1)^2 = -I_4$ and $(\gamma^3)^2 = -I_4$, we can infer the following relations

$$\gamma^1\gamma^3\gamma^0 = \gamma^0(\gamma^1\gamma^3), \quad \gamma^1\gamma^3\gamma^1 = -\gamma^1(\gamma^1\gamma^3), \quad \gamma^1\gamma^3\gamma^2 = \gamma^2(\gamma^1\gamma^3), \quad \gamma^1\gamma^3\gamma^3 = -\gamma^3(\gamma^1\gamma^3). \tag{12.574}$$

A substitution of these expressions into (12.573) leads to

$$-\frac{i\hbar}{c}\gamma^0\frac{\partial\left(\gamma^1\gamma^3\psi^*(t, \vec{r})\right)}{\partial t} = \left(-i\hbar\gamma^1\frac{\partial}{\partial x} - i\hbar\gamma^2\frac{\partial}{\partial y} - i\hbar\gamma^3\frac{\partial}{\partial z} + mc\right)\left(\gamma^1\gamma^3\psi^*(t, \vec{r})\right). \tag{12.575}$$

Since $x'^\mu \equiv (ct', \vec{r}')$ is the time reversal transform of $x^\mu \equiv (ct, \vec{r})$, we have $t = -t'$ and $\vec{r} = \vec{r}'$ and, hence, can write

$$\frac{\partial}{\partial t} = -\frac{\partial}{\partial t'} \qquad \frac{\partial}{\partial x} = \frac{\partial}{\partial x'} \qquad \frac{\partial}{\partial y} = \frac{\partial}{\partial y'} \qquad \frac{\partial}{\partial z} = \frac{\partial}{\partial z'}. \tag{12.576}$$

Now, inserting (12.576) into (12.575), we obtain

$$\frac{i\hbar}{c}\gamma'^0\frac{\partial\left(\gamma^1\gamma^3\psi^*(t,\vec{r})\right)}{\partial t'}=\left(-i\hbar\gamma'^1\frac{\partial}{\partial x'}-i\hbar\gamma'^2\frac{\partial}{\partial y'}-i\hbar\gamma'^3\frac{\partial}{\partial z'}+mc\right)\left(\gamma^1\gamma^3\psi^*(t,\vec{r})\right), \tag{12.577}$$

or

$$i\hbar\frac{\gamma^0}{c}\frac{\partial\psi_T(x'^\mu)}{\partial t'}=\left(-i\hbar\vec{\gamma}\cdot\vec{\nabla}'+mc\right)\psi_T(x'^\mu), \tag{12.578}$$

where $\vec{\nabla}'\equiv(\partial/\partial x',\,\partial/\partial y',\,\partial/\partial z')$ and $\psi_T(x'^\mu)=\gamma^1\gamma^3\psi^*(x^\mu)$.

Comparing 12.569) and 12.578), we see that the Dirac equation is covariant under time reversal provided that the time reversal operator is defined by $\hat{T}=\eta\gamma^1\gamma^3\hat{K}$, where η is a phase factor which, in the case of fermions, can be taken to be $\eta=i$, and where $\hat{K}$ is an operator that simply takes the complex conjugate of the quantities on its right-hand side; for instance: $\hat{K}a=a^*\hat{K}$ (where a is any complex number) and $\hat{K}\psi(t,\vec{r})=\psi^*(t,\vec{r})$. The conjugation operator $\hat{K}$ is defined by its action on kets, not on bras. As mentioned above, $\hat{K}$ also satisfies this condition: $\hat{K}=\hat{K}^\dagger=\hat{K}^{-1}$.

So, the Dirac equation is covariant under time reversal provided that the time reversal operator is given by

$$\boxed{\hat{T}=i\gamma^1\gamma^3\hat{K},} \tag{12.579}$$

and, hence, the Dirac spinors would transform as follows

$$\boxed{\psi_T(t',\vec{r})=\hat{T}\psi(t,\vec{r})=i\gamma^1\gamma^3\hat{K}\,\psi(t,\vec{r})=i\gamma^1\gamma^3\,\psi^*(t,\vec{r})\quad\text{or}\quad\psi_T(-t,\vec{r})=i\gamma^1\gamma^3\psi^*(t,\vec{r}).} \tag{12.580}$$

We may mention that the product of the matrices of γ^1 and γ^3 (see (12.222) and (12.223)) yields the matrix form of the time reversal operator:

$$\hat{T}=i\gamma^1\gamma^3\hat{K}=\begin{pmatrix}0&i&0&0\\-i&0&0&0\\0&0&0&i\\0&0&-i&0\end{pmatrix}\hat{K}. \tag{12.581}$$

We can verify that the matrix (12.581) is equal to its own inverse, which is expected as the time reversal operator is antiunitary: $\hat{T}^\dagger=\hat{T}^{-1}=\hat{T}$. This is also expected as two successive time reversals must take us back to the original, time forward state:

$$\hat{T}^2\psi(t,\vec{r})=\hat{T}\psi_T(-t,\vec{r})=\psi(t,\vec{r})\quad\text{or}\quad\hat{T}^2\equiv\begin{pmatrix}0&i&0&0\\-i&0&0&0\\0&0&0&i\\0&0&-i&0\end{pmatrix}^2=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}=I_4. \tag{12.582}$$

Action of the time reversal operator on the free-particle spinors

By analogy to the calculations we carried out above with the charge conjugation operator (see equations (12.508)–(12.510)), let us see how the free-particle spinors are transformed under time reversal. For this, we are going to consider the case of a free particle moving along

the $z-$axis; the spinors $u_1(p_z)$, $u_2(p_z)$, $v_1(p_z)$, $v_2(p_z)$ can be obtained by setting $p_x = p_y = 0$ in (12.336):

$$u_1(p_z) = N\begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \\ 0 \end{pmatrix}, \quad u_2(p_z) = N\begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix}, \quad v_1(p_z) = N\begin{pmatrix} 0 \\ \frac{-cp_z}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix}, \quad v_2(p_z) = N\begin{pmatrix} \frac{cp_z}{E^{(+)}+mc^2} \\ 0 \\ 1 \\ 0 \end{pmatrix}. \tag{12.583}$$

So applying $\hat{T}$ on the spinors $u_1(p_z)$, $u_2(p_z)$, $v_1(p_z)$, $v_2(p_z)$, which reduces to applying (12.581) on (12.583), we obtain

$$\hat{T}u_1(p_z)^{\uparrow} = -iu_2(-p_z)^{\downarrow}, \qquad \hat{T}u_2(p_z)^{\downarrow} = iu_1(-p_z)^{\uparrow}, \tag{12.584}$$

$$\hat{T}v_1(p_z)^{\uparrow} = iv_2(-p_z)^{\downarrow}, \qquad \hat{T}v_2(p_z)^{\downarrow} = -iv_1(-p_z)^{\uparrow}, \tag{12.585}$$

where the up-arrow ($\uparrow$) stands for an upward spin, and ($\downarrow$) for a downward spin and the i and $-i$ factors are nothing but phase factors which will cancel out when we calculate mean values of observables such as energy. As an illustration, here is how we got the first relation above; namely, $\hat{T}u_1(p_z)^{\uparrow} = -iu_2(-p_z)^{\downarrow}$:

$$\begin{aligned} \hat{T}u_1(p_z) &= \begin{pmatrix} 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix}\hat{K}u_1(p_z) = N\begin{pmatrix} 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \\ 0 \end{pmatrix}^* \\ &= N\begin{pmatrix} 0 \\ -i \\ 0 \\ \frac{-icp_z}{E^{(+)}+mc^2} \end{pmatrix} = -iN\begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \end{pmatrix} \\ &\equiv -iu_2(-p_z). \end{aligned} \tag{12.586}$$

Equations (12.584)–(12.585) shows that the time reversal operator transforms a spin-up particle, with momentum p_z, into a spin-down particle with opposite momentum, $-p_z$, and vice versa; and a spin-up antiparticle, with momentum p_z, into a spin-down antiparticle with opposite momentum, $-p_z$, and vice versa.

In sum, we have seen in this section that the Dirac equation is covariant under the three discrete symmetries of charge conjugation, parity, and time reversal. Table 12.4 provides a synthesis of the main properties of these three discrete symmetries.

Example 12.11
Show how the four-current-density j^μ transforms under time-reversal operations.

Solution

According to the standard time reversal operation, some quantities flip sign, others do not. For instance, some quantities that flip sign are: time, t, linear momentum, $\vec{p}$, electric current, $\vec{j}$, angular momentum, $\vec{J}$, magnetic field, $\vec{B}$, etc:

$$t \xrightarrow{T} t' = -t, \qquad \vec{p} \xrightarrow{T} \vec{p}\,' = -\vec{p}, \qquad \vec{j} \xrightarrow{T} \vec{j}\,' = -\vec{j}. \tag{12.587}$$

Examples of quantities that are invariant under time reversal, one might mention: position vector, $\vec{r}$, electric field, $\vec{E}$, nabla vector, $\vec{\nabla}$, charge densit, ρ, etc:

$$\vec{r} \xrightarrow{T} \vec{r}\,' = \vec{r}, \qquad \vec{\nabla} \xrightarrow{T} \vec{\nabla}\,' = \vec{\nabla}, \qquad \rho \xrightarrow{T} \rho\,' = \rho. \tag{12.588}$$

Hence, since the 4-current density j^μ is given by $j^\mu = (c\rho,\ \vec{j})$ and since the charge density is invariant, while the current density flips sign, we have

$$j^\mu = (c\rho,\ \vec{j}) \xrightarrow{T} j'^\mu = (c\rho,\ -\vec{j}). \tag{12.589}$$

12.4.25 The *CPT* Theorem

In nature, there are processes that violate one or more discrete symmetries; for instance, and as seen above, both parity and charge conjugation are violated in β–decay. However, the discrete transformation *CPT*, which consists of the product of charge conjugation *C*, parity reflection *P*, and time reversal *T* is a symmetry that is true for any relativistic theory which is invariant under Lorentz transformations. This is known as the *CPT theorem*[76]. It states that, even if the *C*, *P*, and *T* are not separately conserved, the combined operation of $\hat{C}$, $\hat{\mathcal{P}}$, and $\hat{T}$, in any order, is an exact symmetry of all interactions and that all physical laws must be invariant under the combined action of these three transformations. That is, if any particle is replaced with its corresponding antiparticle, and the space coordinate and time are reversed, the physical laws are unchanged. The *CPT* theorem is supported by experiment; there is no experimental evidence for *CPT* violation. The *CPT* theorem is one of the most basic assertions of physics.

Now, let us find out how the wave function transforms under *CPT*. Since

$$(t,\ \vec{r}) \xrightarrow{\hat{C}\hat{\mathcal{P}}\hat{T}} (-t,\ -\vec{r}), \qquad \vec{p} \xrightarrow{\hat{C}\hat{\mathcal{P}}\hat{T}} -\vec{p}, \qquad q \xrightarrow{\hat{C}\hat{\mathcal{P}}\hat{T}} -q, \tag{12.590}$$

we can write

$$\psi(t,\ \vec{r}) \xrightarrow{\hat{C}\hat{\mathcal{P}}\hat{T}} \psi_{CPT}(-t,\ -\vec{r}) = \hat{C}\hat{\mathcal{P}}\hat{T}\psi(t,\ \vec{r}). \tag{12.591}$$

Using the definitions of $\hat{C}$, $\hat{\mathcal{P}}$, and $\hat{T}$ from Table 12.4, we can show that the combined action of the $\hat{C}\hat{\mathcal{P}}\hat{T}$ operators on $\psi(t,\ \vec{r})$ leads to

$$\begin{aligned}
\psi_{CPT}(-t,\ -\vec{r}) &= \hat{C}\hat{\mathcal{P}}\hat{T}\psi(t,\ \vec{r}) = \hat{C}\hat{\mathcal{P}}\left(i\gamma^1\gamma^3\psi^*(t,\ \vec{r})\right) \\
&= \hat{C}\left(i\gamma^0\gamma^1\gamma^3\psi^*(t,\ \vec{r})\right) = -i\hat{C}\left(\gamma^0\gamma^1\gamma^3\psi^*(t,\ \vec{r})\right) \\
&= -i^2\gamma^2\gamma^0\gamma^1\gamma^3\psi(t,\ \vec{r}) = \gamma^0\gamma^1\gamma^2\gamma^3\psi(t,\ \vec{r}) \\
&= -i\gamma^5\psi(t,\ \vec{r}),
\end{aligned} \tag{12.592}$$

---

[76] J. Schwinger, Phys. Rev. **82** (6), 914—927 (1951).
G. Lüders, *"On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories"*, K. Dan. Vidensk. Selsk. Mat.-fys. Medd. **28** (5), 1—17 (1954).
W. Pauli: *Exclusion Principle, Lorentz Group and Reflection of Space-Time and Charge*, In Niels Bohr and the Development of Physics, ed. by W. Pauli (Pergamon Press), New York, 30–51 (1955).

Table 12.4: Charge Conjugation , Parity, Time Reversal, and *CPT* symmetries

Symmetry	Transformation	Operator	Wave function transformation
Charge Conjugation	$q \xrightarrow{\hat{C}} -q$ $(t, \vec{r}) \xrightarrow{\hat{C}} (t, \vec{r})$	$\hat{C} = i\gamma^2\hat{K}$	$\psi_C(t, \vec{r}) = \hat{C}\psi(t, \vec{r}) = i\gamma^2\psi^*(t, \vec{r})$
Parity	$(t, \vec{r}) \xrightarrow{\hat{P}} (t, -\vec{r})$	$\hat{P} = \pm\gamma^0$	$\psi_P(t, -\vec{r}) = \hat{P}\psi(t, \vec{r}) = \gamma^0\psi(t, \vec{r})$
Time Reveral	$(t, \vec{r}) \xrightarrow{\hat{T}} (-t, \vec{r})$ $\vec{p} \xrightarrow{\hat{T}} -\vec{p}$	$\hat{T} = i\gamma^1\gamma^3\hat{K}$	$\psi_T(-t, \vec{r}) = \hat{T}\psi(t, \vec{r}) = i\gamma^1\gamma^3\psi^*(t, \vec{r})$
CPT	$(t, \vec{r}) \xrightarrow{\hat{C}\hat{P}\hat{T}} (-t, -\vec{r})$ $\vec{p}, q \xrightarrow{\hat{C}\hat{P}\hat{T}} -\vec{p}, -q$	$\hat{C}\hat{P}\hat{T}$	$\psi_{CPT}(-t, -\vec{r}) = \hat{C}\hat{P}\hat{T}\psi(t, \vec{r}) = -i\gamma^5\psi(t, \vec{r})$

where we have used the fact that $\gamma^2\gamma^0\gamma^1\gamma^3 = -\gamma^0\gamma^2\gamma^1\gamma^3 = \gamma^0\gamma^1\gamma^2\gamma^3$ in the third line. Equation (12.592) is quite significant: if $\psi(t, \vec{r})$ describes a particle traveling forward in space-time, $\psi_{CPT}(-t, -\vec{r})$ describes an antiparticle traveling backwards in space-time; the term $-i\gamma^5$ is merely a multiplicative factor. This result provides an underlying justification for the Feynman-Stükleberg interpretation which states that, in order to preserve causality, positive energy states must propagate forwards in time and negative energy states backwards in time.

The *CPT* theorem leads to a number of experimentally supported conclusions:

- Particles and their corresponding antiparticles have identical masses; for instance, for a neutral kaon particle K^0 and its antiparticle $\overline{K}^0$, we have

$$\frac{M_{K^0} - M_{\overline{K}^0}}{M_{K^0} + M_{\overline{K}^0}} < 10^{-19}. \tag{12.593}$$

- Particles and their corresponding antiparticles have identical lifetimes:

$$\frac{\tau_{K^0} - \tau_{\overline{K}^0}}{\tau_{K^0} + \tau_{\overline{K}^0}} < 10^{-18}. \tag{12.594}$$

- Particles and their corresponding antiparticles have equal, but opposite charges:

$$Q_p + Q_{\overline{p}} < 10^{-21}e. \tag{12.595}$$

- Particles and their corresponding antiparticles have equal magnetic moments; for instance, for an electron and a positron, we have:

$$\frac{\mu_{e^+} - \mu_{e^-}}{\mu_{e^+} + \mu_{e^-}} < 10^{-12}. \tag{12.596}$$

Table 12.5: Conservation laws for the three interactions: *IQN* stands for the Internal Quantum Numbers, *Q* electric charge, *S* Spin, *P* Parity, *C* Charge conjugation, *T* time reversal, *CP* and *CPT* are combined symmetries.

Interaction Type	*IQN*	*Q*	*S*	*P*	*C*	*T*	*CP*	*CPT*
Electromagnetic Interaction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weak Interaction	Yes	Yes	Yes	No	No	No	No	Yes
Strong Interaction	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

- The internal quantum numbers (IQN) of particles are opposite to those of their corresponding antiparticles.

Discrete symmetry breaking in our universe

Taken separately, the three discrete symmetries (*C*, *P*, *T*) are violated by the weak interaction. Additionally, the C and *P* are conserved separately in the strong and electromagnetic interactions as displayed in Table 12.5. Although both *P* and C were established to be violated in weak interactions around 1957, the *CP*−symmetry (the combined actions of *C* and *P*) was thought to be invariant for a number of years. While the electromagnetic and strong interactions seemed to be invariant under the combined *CP*−symmetry, it was established experimentally in 1964 that *CP* is violated in neutral kaon decays[77]. This experiment has demonstrated that the electrically neutral K meson, which was thought to break down into three pi mesons, decayed a fraction of the time into only two such particles (i.e., $K^0 \longrightarrow \pi^+ + \pi^-$), thereby violating *CP*−symmetry. Moreover, the *CP* violation implied nonconservation of the time reversal symmetry *T*, provided that the long-held *CPT* theorem is valid; as a combination, the *CPT* symmetry constitutes an exact symmetry of all types of fundamental interactions.

We should note that, *due to the CP symmetry breaking, the universe is dominated by matter over antimatter; although charge reversal symmetry requires equal amounts of matter and anti-matter, our universe consists mostly of matter.* Additionally, if the *CP*−symmetry is violated, then the *T* symmetry is violated and the physical laws are not invariant under the reversal of the direction of time. Due to *CPT* conservation, the following symmetries are equivalent::

$$CP \leftrightarrow T, \qquad CT \leftrightarrow P, \qquad PT \leftrightarrow C. \tag{12.597}$$

12.5 Successes and Limitations of Relativistic Quantum Mechanics

For almost a century now, relativistic quantum mechanics have been shaped by the Dirac equation. When one talks about relativistic quantum mechanics, one almost always refers to the Dirac equation; the Klein–Gordon plays a junior role in comparison to the Dirac equation. Hence, the successes and shortcomings of relativistic quantum mechanics boil down to those of the Dirac equation.

First, we may summarize the main successes of the Dirac equation as follows:

[77] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Phys. Rev. Let.* 3, 138 (1964).

- Unlike the Pauli-Shrödinger equation which deals with spin in an ad hoc manner, the Dirac equation treats it in a natural way from the very outset; one of its great successes is the fact that its solutions — the 4-component spinors — naturally describe the properties of intrinsic spin.
- The Dirac equation gives the correct value[78] $g \simeq 2$ for the electron gyromagnetic factor, as observed experimentally.
- Unlike the Klein–Gordon equation which gives rise to negative probability densities, the Dirac equation yields probability densities which are always positive.
- The Dirac equation correctly describes the fine structure of the hydrogen atom, while the Schödinger equation deals with it in an *ad hoc* manner by means of perturbation theory.
- The Dirac equation reduces to the Pauli-Schrödinger equation in the non-relativistic limit and, hence, establishing connection with the vast successes of the Shrödinger equation in the realm of non-relativistic quantum mechanics.
- One of the most far-reaching successes of the Dirac equation is its predictive power. It predicted the existence of a new form of matter, *antimatter*, previously unobserved; the existence of the positron was confirmed experimentally several years later. The existence of antimatter is now recognized as a general law of nature: *to every particle, there corresponds an antiparticle*[79]: $e^- \longrightarrow e^+$, $\pi^- \longrightarrow \pi^+$, $p \longrightarrow \overline{p}$, etc.

Second, in spite of its success in providing an accurate description of relativistic electrons, the Dirac equation suffers from a number of shortcomings; chiefly among them are the following interconnected problems:

- Although the Dirac equation started as a one-particle equation, it unwittingly ended up dealing with multiparticle phenomena such as the creation and annihilation of electron-positron pairs. The hole theory is inherently a multiparticle prescription, since it is based on the postulate of an infinity of states that are occupied by negative energy electrons.
- The hole theory, which was introduced by Dirac in an *ad hoc* manner, is conceptually unsatisfactory: the infinite negative charge and infinite negative energy of the Dirac sea are not physically observable; the very idea of an infinite sea of negative energy states is conceptually hard to grasp. The Dirac equation fails to account for these sticky issues and ends up dismissing them altogether with no justification.
- Although the hole theory has enabled us to have an idea about the creation and annihilation of electron-positron pairs in some processes such as the two illustrated in Eqs. (12.326) and (12.327), the Dirac equation cannot describe other processes where single electrons and positrons are created and annihilated. The ability to create and annihilate particles cannot be addressed within a *single-particle* theory such as the Dirac equation.
- The Dirac equation works for fermions only; *it does not work for bosons*. For instance, it provides no explanation for the negative energy solutions of the Klein–Gordon equation.

[78] The calculated value of g matches the experimental value to about 10 significant figures, $g = 2.0023193044$.

[79] Note: some particles, called *self-conjugate*, are identical to their own antiparticles such as neutral pions, photons, etc: $\pi^0 \longrightarrow \pi^0$, $\gamma \longrightarrow \gamma$, etc.

In view of the untractable problems of the Klein–Gordon equation and the above mentioned limitations of the Dirac equation, it becomes clear that these two equations have given us ample evidence that *relativistic quantum mechanical phenomena cannot be described by means of a single-particle theory*. One has to go beyond these two equations.

12.6 Beyond Relativistic Quantum Mechanics

As seen above, relativistic quantum mechanics, as a single particle theory, is plagued by various inconsistencies: negative energy states and negative probability densities. To address the limitations of the Klein–Gordon and the Dirac equations, one has to go beyond them by invoking a theory that must be *multiparticle* in nature. The way out is provided by *Quantum field theory* (QFT). Having resulted *from the combination of classical field theory, special relativity, and quantum mechanics*, QFT offers the proper framework for dealing with relativistic quantum phenomena and multi-particle systems; QFT is much more general.

Since QFT is beyond the scope of this text, we are going to offer a brief introduction to classical field theory in the following chapter. In particular, we will focus on how to derive the Klein–Gordon and Dirac equations using classical fields. We will use a real scalar field to derive the Klein–Gordon equation and a spinor spin-1/2 field to obtain the Dirac equation.

12.7 Solved Problems

Problem 12.1

Consider two inertial frames F and F' where F' is in uniform motion along the x-direction with constant speed $v = 2.7 \times 10^8$ m/s relative to F. Consider also two events that take place in F and whose Minkowski spacetime coordinates are given by the following 4-vectors: $A^\mu = (4, 2, -1, 3)$, $B^\mu = (-3, -1, 1, 2)$.

(a) Find the Lorentz transformation corresponding to this $x-$axis boost.

(b) Find the coordinates A'^μ and B'^μ of these two events as measured by an observer in F'.

(c) Show quantitatively that the squares of A^μ and B^μ are Lorentz invariant.

(d) Show that this Lorentz transformation preserves the causal character (timelike, spacelike, or null) of A^μ and B^μ.

(e) Show quantitatively that the inner product between A^μ and B_μ is a Lorentz invariant.

Solution

(a) Since the speed of F' is $v = 2.7 \times 10^8$ m/s $\equiv 0.9$c (where c is the speed of light) along the $x-$axis, we can obtain at once the values of the rapidity parameter and the Lorentz factor

$$\beta = v/c = 0.9, \qquad \gamma = 1/\sqrt{1-\beta^2} = 2.2942, \qquad \beta\gamma = 2.0646, \tag{12.598}$$

the corresponding Lorentz transformation can be obtained from Equation (12.9):

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2.2942 & -2.0646 & 0 & 0 \\ -2.0646 & 2.2942 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{12.599}$$

(b) Using the Lorentz matrix (12.599), we can obtain the coordinates of the two events in F' as follows:

$$A' = \Lambda A = \begin{pmatrix} 2.2942 & -2.0646 & 0 & 0 \\ -2.0646 & 2.2942 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 5.0476 \\ -3.67 \\ -1 \\ 3 \end{pmatrix}, \tag{12.600}$$

$$B' = \Lambda B = \begin{pmatrix} 2.2942 & -2.0646 & 0 & 0 \\ -2.0646 & 2.2942 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ -1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -4.818 \\ 3.8996 \\ 1 \\ 2 \end{pmatrix}. \tag{12.601}$$

(c) The square of the 4-vector A^μ can be calculated either formally according to

$$A^2 = A_\mu A^\mu = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2 = (4)^2 - (2)^2 - (-1)^2 - (3)^2 = 2, \tag{12.602}$$

or through a matrix product as

$$A^2 = A_\mu A^\mu = \begin{pmatrix} 4 & -2 & 1 & -3 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \\ -1 \\ 3 \end{pmatrix} = (4)(4) + (-2)(2) + (1)(-1) + (-3)(3) = 2. \tag{12.603}$$

Using the first method, we can obtain the square of A'^μ

$$A'^2 = A'_\mu A'^\mu = (5.0476)^2 - (-3.67)^2 - (1)^2 - (-3)^2 = 2.009 \simeq 2. \tag{12.604}$$

Comparing (12.603) and (12.604), we see that $A^2 = A'^2$ and, hence, the square of A^μ is a Lorentz invariant.

Following the same method, we see that B^2 is also a Lorentz invariant as evidenced by $B^2 = B'^2 = 3$:

$$B^2 = (-3)^2 - (1)^2 - (-1)^2 - (-2)^2 = 3, \tag{12.605}$$

$$B'^2 = (-4.818)^2 - (-3.8996)^2 - (-1)^2 - (-2)^2 = 3.006 \simeq 3 = B^2. \tag{12.606}$$

(d) As shown in equations (12.603) and (12.604, since $A^2 > 0$ and $A'^2 > 0$, both of A^μ and A'^μ are timelike; both of them are causal vectors. Hence this Lorentz transformation does conserve the causal nature of vector A^μ. This Lorentz transformation does also conserve the causal nature of vector B^μ, since $B^2 > 0$ and $B'^2 > 0$ as shown in equation (12.606); both of B^μ and B'^μ are timelike vectors.

(e) Using the inner product relation: $A^\mu B_\mu = A^0 B^0 - A^1 B^1 - A^2 B^2 - A^3 B^3$, we have

$$A^\mu B_\mu = (4)(-3) - (2)(-1) - (-1)(1) - (3)(2) = -15, \tag{12.607}$$

$$A'^\mu B'_\mu = (5.0476)(-4.818) - (-3.67)(3.8996) - (-1)(1) - (3)(2) = -15.007 \simeq A^\mu B_\mu, \tag{12.608}$$

which proves the Lorentz invariance of the inner product $A^\mu B_\mu = A'^\mu B'_\mu = -15$.

Problem 12.2

Consider two inertial frames F and F' where F' is in uniform motion along the x-direction with constant speed $v = 2.5 \times 10^8$ m/s relative to F. Consider also two events that take place in F and whose Minkowski spacetime coordinates are given by the following 4-vectors: $A^\mu = (-6, 4, -4, 2)$, $B^\mu = (3, -2, 2, 1)$.

(a) Find the matrix of the Lorentz transformation corresponding to this x–axis boost.

(b) Find the coordinates A'^μ and B'^μ of these two events as measured by an observer in F'.

(c) Show that this Lorentz transformation preserves the causal nature (timelike, spacelike, or null) of A^μ and B^μ.

(d) Show quantitatively that the distance between these two events is a Lorentz invariant.

Solution

(a) Since the speed of F' is $v = 2.5 \times 10^8$ m/s $\equiv 0.833$c (where c is the speed of light) along the x–axis, we can obtain at once the values of the rapidity parameter and the Lorentz factor

$$\beta = v/c = 0.833, \qquad \gamma = 1/\sqrt{1-\beta^2} = 1.807, \qquad \beta\gamma = 1.506, \tag{12.609}$$

the corresponding Lorentz transformation can be obtained from Equation (12.9):

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1.807 & -1.506 & 0 & 0 \\ -1.506 & 1.807 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{12.610}$$

(b) Using the Lorentz matrix (12.610), we can obtain the coordinates of the two events in F' as follows:

$$A' = \Lambda A = \begin{pmatrix} 1.807 & -1.506 & 0 & 0 \\ -1.506 & 1.807 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -6 \\ 4 \\ -4 \\ 2 \end{pmatrix} = \begin{pmatrix} -16.866 \\ 16.264 \\ -4 \\ 2 \end{pmatrix}, \tag{12.611}$$

$$B' = \Lambda B = \begin{pmatrix} 1.807 & -1.506 & 0 & 0 \\ -1.506 & 1.807 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 8.433 \\ -8.132 \\ 2 \\ 1 \end{pmatrix}. \tag{12.612}$$

(c) The squares of the 4-vectors A^μ and A'^μ are given by

$$A^2 = A_\mu A^\mu = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2 = (-6)^2 - (4)^2 - (-4)^2 - (2)^2 = 0, \tag{12.613}$$

$$A'^2 = A'_\mu A'^\mu = (-16.866)^2 - (16.264)^2 - (-4)^2 - (2)^2 = -0.056 \simeq 0. \tag{12.614}$$

Comparing (12.613) and (12.614), we see that A^μ and A'^μ *are both null vectors* and, hence, the Lorentz transformation (12.610) does conserve the causal nature of vector A^μ.

Following the same method, we see that the vector B^μ and its Lorentz transform B'^μ *are also both null vectors* as evidenced by $B^2 = B'^2 = 0$:

$$B^2 = B_\mu B^\mu = (3)^2 - (-2)^2 - (2)^2 - (1)^2 = 0, \tag{12.615}$$

$$B'^2 = B'_\mu B'^\mu = (8.433)^2 - (-8.132)^2 - (2)^2 - (1)^2 = -0.056 \simeq 0 = B^2. \tag{12.616}$$

(d) The vectors separating the two events $A^\mu = (-6,\ 4,\ -4,\ 2)$, $B^\mu = (3,\ -2,\ 2,\ 1)$ and their respective Lorentz transforms $A'^\mu = (-16.866,\ 16.264,\ -4,\ 2)$ and $B'^\mu = (8.433,\ -8.132,\ 2,\ 1)$ are given by

$$\begin{aligned} A^\mu - B^\mu &= \left(A^0 - B^0,\ A^1 - B^1,\ A^2 - B^2,\ A^3 - B^3\right) \\ &= (-9,\ 6,\ -6,\ 1), \qquad (12.617) \\ A'^\mu - B'^\mu &= \left(A'^0 - B'^0,\ A'^1 - B'^1,\ A'^2 - B'2,\ A'^3 - B'^3\right) \\ &= (-25.299,\ 24.396,\ -6,\ 1). \qquad (12.618) \end{aligned}$$

Using these expressions, we can find the distances separating the two events as measured by observers in the inertial frames F and F', respectively:

$$\begin{aligned} (A^\mu - B^\mu)^2 &= \left(A^0 - B^0\right)^2 - \left(A^1 - B^1\right)^2 - \left(A^2 - B^2\right)^2 - \left(A^3 - B^3\right)^2 \\ &= (-9)^2 - (6)^2 - (-6)^2 - (1)^2 = 8, \qquad (12.619) \\ (A'^\mu - B'^\mu)^2 &= \left(A'^0 - B'^0\right)^2 - \left(A'^1 - B'^1\right)^2 - \left(A'^2 - B'^2\right)^2 - \left(A'^3 - B'^3\right)^2 \\ &= (-25.299)^2 - (24.264)^2 - (-6)^2 - (1)^2 = 7.875 \simeq 8, \qquad (12.620) \end{aligned}$$

From equations (12.619) and (12.620), we see that the distance between these two events is a Lorentz invariant; it has the same value of 8 in the inertial frames F and F'.

Problem 12.3
Consider two inertial frames of reference F and F' where F' is in uniform motion along the $y-$axis with constant velocity $\vec{v}$ relative to F. An event is described by the 4-coordinates $(ct,\ x,\ y, z)$ in F and by $(ct',\ x',\ y',\ z')$ in F'. Find the Lorentz transformation that connects $(ct,\ x,\ y, z)$ to $(ct',\ x',\ y',\ z')$.

Solution:
This transformation affects only time and y'; the components x' and z' will not be affected. From the theory of special relativity, we can write:

$$ct' = \gamma(ct - \beta y), \quad x' = x, \quad y' = \gamma(y - c\beta t), \quad z' = z, \qquad (12.621)$$

where $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$. We can rewrite this transformation in the following matrix form:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & -\gamma\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\gamma\beta & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}. \qquad (12.622)$$

Problem 12.4
Consider the following two consecutive boosts: a v_x boost along the $x-$direction followed by a v_y boost along the $y-$direction.

(a) Find the Lorentz transformations corresponding to the two boosts described above.

(b) Show that when the two boosts described above are performed in the reverse order, they would give a different transformation; that is, show that the Lorentz transformation

corresponding to a v_y boost in the y–direction followed by a v_x boost in the x–direction would be different from the result obtained in (a).

(c) Illustrate numerically the results of (a) and (b) on the case where $v_x = 2.7 \times 10^6$ m/s and $v_y = 2.5 \times 10^8$ m/s.

Solution (a) The Lorentz transformations corresponding to the v_x and v_y boosts can be inferred at once from (12.9) and (12.622)

$$[\Lambda_x] = \begin{pmatrix} \gamma_x & -\beta_x\gamma_x & 0 & 0 \\ -\beta_x\gamma_x & \gamma_x & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{12.623}$$

$$]\Lambda_y] = \begin{pmatrix} \gamma_y & 0 & -\beta_y\gamma_y & 0 \\ 0 & 1 & 0 & 0 \\ -\beta_y\gamma_y & 0 & \gamma_y & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{12.624}$$

where

$$\beta_x = v_x/c, \qquad \gamma_x = 1/\sqrt{1-\beta_x^2}, \tag{12.625}$$

$$\beta_y = v_y/c, \qquad \gamma_y = 1/\sqrt{1-\beta_y^2}. \tag{12.626}$$

Hence, the Lorentz transformation corresponding to a v_x boost followed by a v_y is given by the product of (12.623) and (12.624)

$$\begin{aligned} [\Lambda_x][\Lambda_y] &= \begin{pmatrix} \gamma_x & -\beta_x\gamma_x & 0 & 0 \\ -\beta_x\gamma_x & \gamma_x & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \gamma_y & 0 & -\beta_y\gamma_y & 0 \\ 0 & 1 & 0 & 0 \\ -\beta_y\gamma_y & 0 & \gamma_y & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \gamma_x\gamma_y & -\beta_x\gamma_x & -\beta_y\gamma_x\gamma_y & 0 \\ -\beta_x\gamma_x\gamma_y & \gamma_x & \beta_x\beta_y\gamma_x\gamma_y & 0 \\ -\beta_y\gamma_y & 0 & \gamma_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{12.627}$$

(b) The Lorentz transformation corresponding to a v_y boost followed by a v_x boost can be obtained from the product of (12.624) and (12.623):

$$\begin{aligned} [\Lambda_y][\Lambda]_x &= \begin{pmatrix} \gamma_y & 0 & -\beta_y\gamma_y & 0 \\ 0 & 1 & 0 & 0 \\ -\beta_y\gamma_y & 0 & \gamma_y & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \gamma_x & -\beta_x\gamma_x & 0 & 0 \\ -\beta_x\gamma_x & \gamma_x & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \gamma_x\gamma_y & -\beta_x\gamma_x\gamma_y & -\beta_y\gamma_y & 0 \\ -\beta_x\gamma_x & \gamma_x & 0 & 0 \\ -\beta_y\gamma_x\gamma_y & \beta_x\beta_y\gamma_x\gamma_y & \gamma_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{12.628}$$

As expected, comparing (12.627) and (12.628), we see that the the Lorentz transformation corresponding to a v_x boost followed by a v_y boost is different from a v_y boost followed by a v_x boost.

(c) Using the numerical values

$$\beta_x = v_x/c = 0.9, \qquad \gamma_x = 1/\sqrt{1-\beta_x^2} = 2.2942, \qquad \beta_x\gamma_x = 2.0646, \tag{12.629}$$

$$\beta_y = v_y/c = 0.833, \qquad \gamma_y = 1/\sqrt{1-\beta_y^2} = 1.807, \qquad \beta_y\gamma_y = 1.506, \tag{12.630}$$

we can obtain the Lorentz transformations of the $v_x = 2.7\times10^8$ m/s and the $v_y = 2.5\times10^8$ m/s boosts

$$[\Lambda_x] = \begin{pmatrix} \gamma_x & -\beta_x\gamma_x & 0 & 0 \\ -\beta_x\gamma_x & \gamma_x & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2.2942 & -2.0646 & 0 & 0 \\ -2.0646 & 2.2942 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{12.631}$$

$$\Lambda_y = \begin{pmatrix} \gamma_y & 0 & -\beta_y\gamma_y & 0 \\ 0 & 1 & 0 & 0 \\ -\beta_y\gamma_y & 0 & \gamma_y & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1.807 & 0 & -1.506 & 0 \\ 0 & 1 & 0 & 0 \\ -1.506 & 0 & 1.807 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{12.632}$$

Hence, the Lorentz transformation corresponding to the $v_x = 2.7\times10^8$ m/s boost followed by the $v_y = 2.5\times10^8$ m/s is given by the product of (12.631) and (12.632)

$$\begin{aligned}[\Lambda_x][\Lambda_y] &= \begin{pmatrix} 2.2942 & -2.0646 & 0 & 0 \\ -2.0646 & 2.2942 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1.807 & 0 & -1.506 & 0 \\ 0 & 1 & 0 & 0 \\ -1.506 & 0 & 1.807 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 4.146 & -2.065 & -3.455 & 0 \\ -3.731 & 2.294 & 3.109 & 0 \\ -1.506 & 0 & 1.807 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{12.633}$$

Similarly, the Lorentz transformation corresponding to the $v_y = 2.5\times10^8$ m/s boost followed by the $v_x = 2.7\times10^8$ m/s can be obtained from the product of (12.632) and (12.631):

$$\begin{aligned}[\Lambda_y][\Lambda_x] &= \begin{pmatrix} 1.807 & 0 & -1.506 & 0 \\ 0 & 1 & 0 & 0 \\ -1.506 & 0 & 1.807 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2.2942 & -2.0646 & 0 & 0 \\ -2.0646 & 2.2942 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 4.146 & -3.731 & -1.506 & 0 \\ -2.065 & 2.294 & 0 & 0 \\ -3.455 & 3.109 & 1.87 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{12.634}$$

As expected, comparing (12.633) and (12.634), we see that the the Lorentz transformation corresponding to a $v_x = 2.7\times10^8$ m/s boost followed by a $v_y = 2.5\times10^8$ m/s boost is

different from the transformation corresponding to a $v_y = 2.5 \times 10^8$ m/s boost followed by a $v_x = 2.7 \times 10^8$ m/s boost.

Problem 12.5

(a) Find the Minkowski metric, g, for a 4-dimensional spacetime in spherical coordinates.
(b) Find the inverse of the metric derived in (a) and verify that $[g]\,[g]^{-1} = I_4$.

Solution

(a) As shown in Appendix D, the spacetime interval between two infinitely neighboring events is given by

$$ds^2 = f_{x^0}\left(dx^0\right)^2 + f_{x^1}\left(dx^1\right)^2 + f_{x^2}\left(dx^2\right)^2 + f_{x^3}\left(dx^3\right)^2, \tag{12.635}$$

where f_{x^0}, f_{x^1}, f_{x^2}, f_{x^3} are scale factors that account for the space's curvature. Since, for a *flat* Minkowski spacetime, these scale factors are given

$$f_{x^0} = 1, \qquad f_{x^1} = f_{x^2} = f_{x^3} = -1, \tag{12.636}$$

the interval expression (12.635) becomes

$$ds^2 = \left(dx^0\right)^2 - \left(dx^1\right)^2 - \left(dx^2\right)^2 - \left(dx^3\right)^2. \tag{12.637}$$

Using the fact that $dx^0 = cdt$, $dx^1 = dx$, $dx^2 = dy$, $dx^3 = dz$, we write (12.637) as

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2. \tag{12.638}$$

Now, in spherical coordinates $x = (x^0, x^1, x^2, x^3) = (ct, r, \theta, \phi)$ where $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $x = r\cos\theta$, we can write

$$\begin{aligned} dx &= \frac{\partial x}{\partial r}dr + \frac{\partial x}{\partial \theta}d\theta + \frac{\partial x}{\partial \phi}d\phi \\ &= \sin\theta\cos\phi\, dr + r\cos\theta\cos\phi\, d\theta - r\sin\theta\sin\phi\, d\theta, \qquad (12.639) \\ dy &= \frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial \theta}d\theta + \frac{\partial y}{\partial \phi}d\phi \\ &= \sin\theta\cos\phi\, dr + r\cos\theta\sin\phi\, d\theta + r\sin\theta\cos\phi\, d\theta, \qquad (12.640) \\ dz &= \frac{\partial z}{\partial r}dr + \frac{\partial z}{\partial \theta}d\theta + \frac{\partial z}{\partial \phi}d\phi \\ &= \cos\theta\, dr - r\sin\theta\, d\theta. \qquad (12.641) \end{aligned}$$

Using these expressions, we can carry out the following calculation straightforwardly

$$\begin{aligned} dx^2 + dy^2 + dz^2 &= \left(dr\sin\theta\cos\phi + r\cos\theta\cos\phi d\theta - r\sin\theta\sin\phi d\theta\right)^2 \\ &\quad + \left(dr\sin\theta\cos\phi + r\cos\theta\sin\phi d\theta + r\sin\theta\cos\phi d\theta\right)^2 \\ &\quad + (dr\cos\theta - r\sin\theta d\theta)^2 \\ &= dr^2 + r^2d\theta^2 + r^2\sin^2 d\phi^2. \end{aligned} \tag{12.642}$$

Inserting (12.642) into (12.638), we obtain

$$\begin{aligned} ds^2 &= c^2dt^2 - dx^2 - dy^2 - dz^2 \\ &= c^2dt^2 - dr^2 - r^2d\theta^2 - r^2\sin^2 d\phi^2. \end{aligned} \tag{12.643}$$

On the other hand, since the infinitesimal interval between two events in spacetime is given by

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu \equiv g_{00}\left(dx^0\right)^2 + g_{11}\left(dx^1\right)^2 + g_{22}\left(dx^2\right)^2 + g_{33}\left(dx^3\right)^2, \tag{12.644}$$

we can rewrite (12.643) as

$$\begin{aligned} ds^2 &= c^2dt^2 - dr^2 - r^2d\theta^2 - r^2\sin^2 d\phi^2 \\ &= g_{00}dt^2 + g_{rr}dr^2 + g_{\theta\theta}d\theta^2 + g_{\phi\phi}d\phi^2, \end{aligned} \tag{12.645}$$

which leads to

$$g_{00} = 1, \quad g_{rr} = -1, \quad g_{\theta\theta} = -r^2, \quad g_{\phi\phi} = -r^2\sin\theta^2, \quad \text{and} \quad g_{\mu\nu} = 0 \quad \text{for} \quad \mu \neq \nu. \tag{12.646}$$

Hence, the Minkowski metric in spherical coordinates is given by

$$[g] = \left(g_{\mu\nu}\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix}. \tag{12.647}$$

(b) We can easily show that the inverse matrix of (12.647) is given by

$$[g]^{-1} = \left(g^{\mu\nu}\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -\frac{1}{r^2} & 0 \\ 0 & 0 & 0 & \frac{1}{r^2\sin^2\theta} \end{pmatrix}. \tag{12.648}$$

We can easily verify that the product of (12.647) and (12.648) yields the identity matrix

$$[g]\,[g]^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -\frac{1}{r^2} & 0 \\ 0 & 0 & 0 & \frac{1}{r^2\sin^2\theta} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = I_4. \tag{12.649}$$

Problem 12.6
At what energies will the following be considered relativistic: (a) photons; (b) electrons; (c) protons?

Solution

(a) Photons are *massless* and, hence, they cannot be non-relativistic; they are always relativistic.

(b) Massive particles are relativistic when their kinetic energy is comparable to or greater than their rest-mass energy m_0c^2. In other words, a massive particle is relativistic when its total mass-energy (rest mass + kinetic energy) is at least twice its rest mass:

$$E \geq 2m_0c^2. \tag{12.650}$$

Since the rest-mass energy of an electron is $m_0c^2 = 0.511$ MeV, it is relativistic only if has sufficient total energy:

$$E \geq 2m_0c^2 = 2 \times 0.511 \text{ MeV} = 1.022 \text{ MeV}. \tag{12.651}$$

(b) Since the rest-mass energy of a proton is $m_0c^2 = 938.27$ MeV, it is relativistic only if its total energy is in the range of:

$$E \geq 2m_0c^2 = 2 \times 938.27 \text{ MeV} = 1.877 \text{ GeV}. \tag{12.652}$$

Problem 12.7
Consider an electron moving with a speed of 2.7×10^8 m/s.
(a) Calculate the kinetic energy of the electron using the relativistic formula.
(b) Calculate the kinetic energy of the electron using the classical formula and compare it with the relativistic result obtained in Part (a).
(c) Estimate the amount of energy needed to accelerate a 1 kg object to a speed of 2.7×10^8 m/s.

Solution
(a) At a speed of $v = 2.7 \times 10^8$ m/s $\equiv 0.9c$, where $c = 3 \times 10^8$ m/s, the electron is relativistic. The relativistic kinetic energy is given by

$$K_{rel} = mc^2 - m_0c^2, \tag{12.653}$$

where m_0 is the rest mass of the electron and m is its mass when moving at a speed v:

$$m = \gamma m_0 = \frac{1}{\sqrt{1 - v^2/c^2}} m_0 \equiv \frac{1}{\sqrt{1 - (0.9c)^2/c^2}} m_0 = 2.294 m_0. \tag{12.654}$$

Using the fact that the rest-mass energy of the electron is given by $m_0c^2 = 0.511$ MeV, and inserting (12.654) into (12.653), we obtain

$$K_{rel} = mc^2 - m_0c^2 = (m - m_0)c^2 = 1.294 m_0c^2 \equiv 0.661 \text{ MeV}. \tag{12.655}$$

(b) The classical kinetic energy can be obtained as follows: when an object moves with a speed much smaller than the speed of light, $v \ll c$, a Taylor expansion in the small parameter v/c leads to

$$m = \gamma m_0 = \frac{1}{\sqrt{1 - v^2/c^2}} m_0 \simeq \left(1 + \frac{v^2}{2c^2}\right) m_0 = m_0 + \frac{m_0 v^2}{2c^2}, \tag{12.656}$$

which allows us to show that the relativistic kinetic energy reduces to the classical expression:

$$K_{rel} = mc^2 - m_0c^2 \simeq m_0c^2 + \frac{1}{2} m_0 v^2 - m_0c^2 = \frac{1}{2} m_0 v^2 \equiv K_{classical}. \tag{12.657}$$

So, the classical kinetic energy of an electron moving at $v = 2.7 \times 10^8$ m/s is given by the standard expression $K_{classical} = m_0 v^2/2$:

$$K_{classical} = \frac{1}{2} m_0 v^2 = \frac{1}{2} m_0 (0.9c)^2 = 0.405 m_0 c^2 = 0.405(0.511 \text{ MeV}) = 0.207 \text{ MeV}. \quad (12.658)$$

This value is almost three times smaller than the relativistic value

$$K_{rel}/K_{classical} = 0.661/0.207 = 3.20. \quad (12.659)$$

This is due to the fact that it takes an enormous amount of energy to accelerate a particle to a speed close to that of light (relativistic speed).

(c) To accelerate a 1 kg object to a speed of 2.7×10^8 m/s, one would need an energy of about

$$K_{classical} = \frac{1}{2} m_0 v^2 = \frac{1}{2} (1 \text{ kg}) \left(2.7 \times 10^8 \text{ m/s}\right)^2 = 3.65 \times 10^{15} \text{ J}. \quad (12.660)$$

This energy is almost equivalent to the yield of a 1 Megaton thermonuclear bomb (1 Megaton of TNT $\simeq 4.184 \times 10^{15} J$)! So, to impart a relativistic speed on a macroscopic object, one would need an enormous amount of energy. Evidently, no massive object can attain the speed of light, since, according to the relativistic mass formula $m = m_0/\sqrt{1 - v^2/c^2}$, the object's mass will increase with its velocity v and will tend to infinity, $m \longrightarrow \infty$, when the velocity approaches the speed of light, $v \longrightarrow c$.

On the other hand, microscopic particles would not require that much energy to accelerate to relativistic speeds. For instance, the 0.661 MeV energy needed to accelerate an electron to 2.7×10^8 m/s is easily achievable in todays' accelerators. For instance, and as mentioned in the Introduction of this chapter, protons can be accelerated in the Large Hadron Collider to speeds of about $0.999999991c$ using a 7 TeV beam of protons (1 TeV = 10^{12} eV).

Problem 12.8
Show that the Lorentz transformation (12.9) conserves the length square of 4-vectors; i.e., show that $c^2t'^2 - (x'^2 + y'^2 + z'^2) = c^2t^2 - (x^2 + y^2 + z^2)$.

Solution
Using (12.9), we can write

$$c^2t'^2 = \gamma^2 (ct - \beta x)^2 = \gamma^2 \left(c^2t^2 - 2c\beta xt + \beta^2 x^2\right), \quad (12.661)$$

$$x'^2 = \gamma^2 (x - c\beta t)^2 = \gamma^2 \left(x^2 - 2c\beta xt + c^2\beta^2 t^2\right), \quad (12.662)$$

$$y'^2 = y^2, \quad (12.663)$$

$$z'^2 = z^2. \quad (12.664)$$

Subtracting (12.662) from (12.661) and using the fact that $\gamma^2 = 1/(1 - \beta^2)$ and $\gamma^2(1 - \beta^2) = 1$, we obtain

$$\begin{aligned} c^2t'^2 - x'^2 &= \gamma^2 \left(c^2t^2 - 2c\beta xt + \beta^2 x^2 - x^2 + 2c\beta xt - c^2\beta^2 t^2\right) \\ &= \gamma^2 \left(c^2t^2 - x^2\right) - \gamma^2\beta^2 \left(c^2t^2 - x^2\right) \\ &= \gamma^2 \left(1 - \beta^2\right)\left(c^2t^2 - x^2\right) \\ &= c^2t^2 - x^2. \end{aligned} \quad (12.665)$$

Now, subtracting (12.663) and (12.664) from (12.665), we obtain

$$c^2t'^2 - (x'^2 + y'^2 + z'^2) = c^2t^2 - (x^2 + y^2 + z^2) \quad \longrightarrow \quad x'^\mu x'_\mu = x^\mu x_\mu, \tag{12.666}$$

which proves the invariance of the magnitude square of 4-vector under Lorentz transformations.

Problem 12.9
Find the solutions of the Klein–Gordon equation for a free particle moving along the $z-$axis with a momentum p.

Solution

For a free particle moving along the $z-$axis, the wave function will only depend on z and time; the action of the momentum operator on the wave function is given by: $\hat{p}_z\psi = -i\hbar\partial\psi/\partial z$. Hence, the Klein–Gordon equation (12.51) yields

$$\left(\partial_\mu\partial^\mu + \frac{m^2c^2}{\hbar^2}\right)\psi(z,\ t) \equiv \left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial z^2} + \frac{m^2c^2}{\hbar^2}\right)\psi(z,\ t) = 0 \tag{12.667}$$

which can be reduced to

$$-\hbar^2\frac{\partial^2\psi(z,\ t)}{\partial t^2} = \left(-\hbar^2c^2\frac{\partial^2}{\partial z^2} + m^2c^4\right)\psi(z,\ t). \tag{12.668}$$

As seen in Chapter 3, we can attempt solutions that are separable; namely, we can write $\psi(z,\ t)$ as a product of time-dependent and space-dependent functions

$$\psi(z,\ t) = f(t)\phi(z). \tag{12.669}$$

Substituting (12.669) into (12.668) and dividing both sides by $f(t)\phi(z)$, we obtain

$$-\hbar^2\frac{1}{f(t)}\frac{d^2f(t)}{dt^2} = \frac{1}{\phi(z)}\left(-\hbar^2c^2\frac{d^2\phi(z)}{dz^2} + m^2c^4\phi(z)\right). \tag{12.670}$$

Since the left-hand side of this equation depends on time only and the right-hand side on z only, both sides must be equal to a constant. From dimensional analysis, we can easily ascertain that this constant has the dimensions of an energy squared, which we are going to denote by E^2. We can then break (12.670) into two separate differential equations, one depending on time only

$$\frac{d^2f(t)}{dt^2} = -\frac{E^2}{\hbar^2}f(t), \tag{12.671}$$

and the other one on the spatial variable z

$$\left(-\hbar^2c^2\frac{d^2}{dz^2} + m^2c^4\right)\phi(z) = E^2\phi(z). \tag{12.672}$$

The solution of (12.671) can be written as

$$f(t) = e^{-iEt/\hbar}. \tag{12.673}$$

Rewriting (12.672) as

$$\left(\frac{d^2}{dz^2} + k^2\right)\phi(z) = 0, \tag{12.674}$$

where

$$k^2 = \frac{E^2 - m^2c^4}{\hbar^2c^2}, \tag{12.675}$$

we can write the most general solution of (12.674) as

$$\phi(z) = Ae^{ikz} + Be^{-ikz}, \tag{12.676}$$

where A and B are unknown constants. The first and second terms of (12.676) represent plane waves traveling along the positive and negative z−axis, respectively. The second term has to be dropped since the particle under consideration is traveling along the positive z−axis

$$\phi(z) = Ae^{ikz}. \tag{12.677}$$

Combining equations (12.669), (12.673) and (12.677), we can write

$$\psi(z,\ t) = Ae^{i(kz-Et/\hbar)} = Ae^{i(pz-Et)/\hbar}, \tag{12.678}$$

where $p = \hbar k$ is the momentum of the free particle as it travels along the positive z−axis. Now, using equation (12.675), we can infer the following expression:

$$E^2 = \hbar^2k^2c^2 + m^2c^4 = p^2c^2 + m^2c^4. \tag{12.679}$$

This relation yields, as expected, two values for the energy; one positive and the other one negative

$$E_1 = E^{(+)} = \sqrt{p^2c^2 + m^2c^4}, \qquad E_2 = E^{(-)} = -\sqrt{p^2c^2 + m^2c^4} \tag{12.680}$$

Hence, taking both of these two energy expressions into consideration, we can write the most general solution corresponding to (12.678) as

$$\psi(z,\ t) = e^{ipz/\hbar}\left(A_1e^{-iE^{(-)}t/\hbar} + A_2e^{-iE^{(+)}t/\hbar}\right), \tag{12.681}$$

where A_1 and A_2 are unknown constants. Notice that the exponents' signs of the first and second terms are reversed. The first and second terms correspond to negative and positive energy solutions, respectively, since

$$i\frac{\partial}{\partial t}e^{-iE^{(-)}t/\hbar} = E^{(-)}e^{-iE^{(-)}t/\hbar}, \qquad \text{and} \qquad i\frac{\partial}{\partial t}e^{-iE^{(+)}t/\hbar} = E^{(+)}e^{-iE^{(+)}t/\hbar}, \tag{12.682}$$

As expected, the Klein–Gordon equation yields a negative energy solution for which there is no satisfactory interpretation within the theory.

Note that, since the Klein–Gordon equation contains second order time derivations, one would need two boundary conditions to determine the constants A_1 and A_2. However, and as discussed in the text, the Klein–Gordon formalism offers no prescription for finding the second boundary condition. This extra degree of freedom is not present in the Schrödinger equation which, due to its first order dependence on the time derivative, requires only one boundary condition to fully specify the wave function. So, knowing the wave function at a time t, the Klein–Gordon formalism provides no prescription for determining the wave function at later times.

Problem 12.10
Show that the Dirac matrices satisfy the following relation:

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4, \qquad \mu,\nu = 0,\ 1,\ 2,\ 3, \tag{12.683}$$

Where [g]is the Minkowski metric; i.e., $g^{00} = 1$ and $g^{11} = g^{22} = g^{33} = -1$ and $g^{\mu\nu} = 0$ when $\mu \neq \nu$ as defined in Eq. (12.18).

Solution
First, let us recall the forms of the four Dirac matrices:

$$\gamma^0 = \beta, \qquad \gamma^1 = \beta\alpha_x, \qquad \gamma^2 = \beta\alpha_y, \qquad \gamma^3 = \beta\alpha_z, \tag{12.684}$$

where the matrices of β and those of the Pauli matrices are listed in Eqs (12.222)–(12.223). Using the relations (12.153)–(12.155), we can write:

$$\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = 1, \tag{12.685}$$

and

$$\alpha_x\alpha_y + \alpha_y\alpha_x = \alpha_y\alpha_z + \alpha_z\alpha_y = \alpha_z\alpha_x + \alpha_x\alpha_z = 0, \tag{12.686}$$

$$\alpha_x\beta + \beta\alpha_x = \alpha_y\beta + \beta\alpha_y = \alpha_z\beta + \beta\alpha_z = 0. \tag{12.687}$$

To prove the relation[80] $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$ in the most general case, let us consider separately the cases where $\mu = \nu$ and $\mu \neq \nu$.
(a) Case $\mu = \nu$
First, when $\mu = \nu = 0$ and using Eqs (12.684)–(12.685) and since $\gamma^0\gamma^0 = \beta^2 = 1$ and $g^{00} = 1$, the relation (12.683) is self-evidently satisfied:

$$\gamma^0\gamma^0 + \gamma^0\gamma^0 = 2g^{00}, \qquad \text{or} \qquad \gamma^0\gamma^0 = 1. \tag{12.688}$$

Second, when $\mu = \nu = i$ (where $i = 1,\ 2,\ 3$) and using Eqs (12.684)–(12.685) and since $g^{ii} = -1$, we can reduce Eq (12.683 to

$$2\gamma^j\gamma^j = -2 \qquad \text{or} \qquad \gamma^j\gamma^j = -1, \qquad j = 1,\ 2,\ 3. \tag{12.689}$$

Since $\gamma^j = \beta\alpha_j$ (see Eq (12.684)), since $\beta\alpha_j = -\alpha_j\beta$ (see Eq (12.687)), and since $\alpha_j^2 = 1$ and $\beta^2 = 1$ (see Eq (12.685)), we can rewrite (12.689) as

$$\gamma^j\gamma^j = \big(\beta\alpha_j\big)\big(\beta\alpha_j\big) = \big(\beta\alpha_j\big)\big(-\alpha_j\beta\big) = -\beta\alpha_j^2\beta = -1 \qquad \text{or} \qquad -\beta^2 = -1 \qquad \text{or} \qquad 1 = 1. \tag{12.690}$$

In summary, Eqs (12.688) and (12.690) prove that the relation $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4$ hods for $\mu = \nu$.
(b) Case $\mu \neq \nu$
When $\mu \neq \nu$, the right-hand side of Eq (12.683) is zero as can be seen from (12.18), since $g^{\mu\nu} = 0$ when $\mu \neq \nu$.

[80]Sometimes, we omit I_4 from certain relations when it becomes too distractive; however, one has to keep in mind that the γ–matrices are 4×4 matrices.

First Case: $\mu = 0$ and $\nu \neq 0$ (i.e., $\nu = 1, 2, 3$):
One the one hand, using (12.684) and then (12.685), we can write $\gamma^\mu\gamma^\nu$ as

$$\gamma^0\gamma^\nu = (\beta)(\beta\alpha_\nu) = \beta^2\alpha_\nu = \alpha_\nu. \tag{12.691}$$

On the other hand, invoking (12.687) and then (12.685), we can write $\gamma^\nu\gamma^\mu$ as

$$\gamma^\nu\gamma^0 = (\beta\alpha_\nu)(\beta) = (-\alpha_\nu\beta)(\beta) = -\alpha_\nu\beta^2 = -\alpha_\nu. \tag{12.692}$$

Adding (12.691) and (12.692), we see that

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 0 \qquad \text{for} \qquad \mu = 0,\ \nu = 1, 2, 3. \tag{12.693}$$

The same relation applies for $\nu = 0$, $\mu = 1, 2, 3$.
Second Case: $\mu \neq \nu$ while $\mu \neq 0$ and $\nu \neq 0$:
One the one hand, using (12.687) and then (12.685), we can write $\gamma^\mu\gamma^\nu$ as

$$\gamma^\mu\gamma^\nu = \left(\beta\alpha_\mu\right)(\beta\alpha_\nu) = \left(-\alpha_\mu\beta\right)(\beta\alpha_\nu) = -\alpha_\mu\beta^2\alpha_\nu = -\alpha_\mu\alpha_\nu. \tag{12.694}$$

Similarly, on the other hand, invoking (12.687) and then (12.685), we can write $\gamma^\nu\gamma^\mu$ as

$$\gamma^\nu\gamma^\mu = (\beta\alpha_\nu)\left(\beta\alpha_\mu\right) = (-\alpha_\nu\beta)\left(\beta\alpha_\mu\right) = -\alpha_\nu\beta^2\alpha_\mu = -\alpha_\nu\alpha_\mu \tag{12.695}$$

Now, adding (12.694) and (12.695), we obtain

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = -\alpha_\mu\alpha_\nu - \alpha_\nu\alpha_\mu = -\left(\alpha_\mu\alpha_\nu + \alpha_\nu\alpha_\mu\right) = 0, \tag{12.696}$$

where we have used Eq (12.686) in the last equation.

In summary, Eqs (12.693) and (12.696) prove that the relation $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4$ hods for $\mu \neq \nu$.

Finally, combining Eqs (12.688), (12.690), Eq(12.693), and (12.696), we see that $\gamma^\mu\gamma^\nu+\gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4$ hods for all contingencies.

Problem 12.11
Using the algebraic properties of the γ–matrices, prove the following relations
(a) Tr $\gamma^5 = 0$.
(b) Tr $(\gamma^\mu) = 0$.
(c) Tr $(\gamma^\mu\gamma^\nu) = 4g_{\mu\nu}$.

Solution:

(a) Using the relation $\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\nu$ for $\mu \neq \nu$, we can pull γ^0 all the way to the right as follows:

$$\begin{aligned} \text{Tr}\,\gamma^5 &= i\text{Tr}\left(\gamma^0\gamma^1\gamma^2\gamma^3\right) = -i\text{Tr}\left(\gamma^1\gamma^0\gamma^2\gamma^3\right) = i\text{Tr}\left(\gamma^1\gamma^2\gamma^0\gamma^3\right) \\ &= -i\text{Tr}\left(\gamma^1\gamma^2\gamma^3\gamma^0\right). \end{aligned} \tag{12.697}$$

Now, since the trace of a product of matrices (or operators) is invariant under their cyclic permutations, for any two matrices A and B, we have

$$\text{Tr}\,(AB) = \text{Tr}\,(BA).$$

Taking $A = \gamma^1\gamma^2\gamma^3$ and $B = \gamma^0$, we can reduce (12.697) to:

$$\text{Tr}\,\gamma^5 = -i\text{Tr}\left[\left(\gamma^1\gamma^2\gamma^3\right)\gamma^0\right] = -i\text{Tr}\left(\gamma^0\gamma^1\gamma^2\gamma^3\right) = -\text{Tr}\,\gamma^5, \tag{12.698}$$

and, hence, we have

$$\text{Tr}\,\gamma^5 = 0. \tag{12.699}$$

(b) Using the facts that $(\gamma^5)^2 = I_4$ and that the trace of a product of matrices is invariant under their cyclic permutations, we can write

$$\text{Tr}\,(\gamma^\mu) = \text{Tr}\left(\gamma^\mu\gamma^5\gamma^5\right) = \text{Tr}\left(\gamma^5\gamma^\mu\gamma^5\right). \tag{12.700}$$

Now, invoking the relation $\gamma^5\gamma^\mu = -\gamma^\mu\gamma^5$, we can move the left γ^5 to the right-hand side of γ^μ and, hence, rewrite (12.700) as

$$\text{Tr}\,(\gamma^\mu) = \text{Tr}\left(\gamma^5\gamma^\mu\gamma^5\right) = -\text{Tr}\left(\gamma^\mu\gamma^5\gamma^5\right) = -\text{Tr}\,(\gamma^\mu) \quad\Longrightarrow\quad \text{Tr}\,(\gamma^\mu) = 0; \tag{12.701}$$

In the last step, we have used again the relation $(\gamma^5)^2 = I_4$.

(c) Using the fact that

$$\text{Tr}\,(\gamma^\mu\gamma^\nu) = \frac{1}{2}\left(\text{Tr}\,(\gamma^\mu\gamma^\nu) + \text{Tr}\,(\gamma^\nu\gamma^\mu)\right), \tag{12.702}$$

we can write

$$\begin{aligned}
\text{Tr}\,(\gamma^\mu\gamma^\nu) &= \frac{1}{2}\left(\text{Tr}\,(\gamma^\mu\gamma^\nu) + \text{Tr}\,(\gamma^\nu\gamma^\mu)\right) \\
&= \frac{1}{2}\text{Tr}\,(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu) \\
&= \frac{1}{2}\text{Tr}\,(2g^{\mu\nu}I_4) \\
&= g^{\mu\nu}\text{Tr}\,(I_4) \\
&= 4g^{\mu\nu},
\end{aligned} \tag{12.703}$$

where we have used the relation $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4$ in the third line and $\text{Tr}\,(I_4) = 4$ in the fifth line.

Problem 12.12

(a) Solve the eigenvalue problem displayed in equations (12.315)–(12.318) for a free-particle and find the eigenstates along with the corresponding energies.

(b) Discuss the normalization conditions of the eigenstates obtained in (a).

Solution

(a) To obtain the solutions of (12.315)–(12.318), we proceed as follows. First, invoking the matrix forms (12.162) of the 2×2 Pauli matrices, we can write the quantity $\vec{\sigma}\cdot\vec{p}$ in the following 2×2 matric form:

$$\begin{aligned}
\vec{\sigma}\cdot\vec{p} &= \sigma_x p_x + \sigma_y p_y + \sigma_z p_z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} p_x + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} p_y + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} p_z \\
&= \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix}.
\end{aligned} \tag{12.704}$$

Let us now write the 4-component spinor (12.307) in terms of the two bi-spinors u_A and u_B as follows

$$u(p^\mu) = N\begin{pmatrix} u_1(p^\mu) \\ u_2(p^\mu) \\ u_3(p^\mu) \\ u_4(p^\mu) \end{pmatrix} = N\begin{pmatrix} u_A \\ u_B \end{pmatrix} \qquad \text{where} \qquad u_A = \begin{pmatrix} u_1(p^\mu) \\ u_2(p^\mu) \end{pmatrix}, \qquad u_B = \begin{pmatrix} u_3(p^\mu) \\ u_4(p^\mu) \end{pmatrix}, \tag{12.705}$$

and where N is a constant (normalization factor which we will deal with below). Combining (12.704) and (12.705) along with (12.315)–(12.318), we can rewrite (12.316)–(12.318) in the following simpler matrix form:

$$\begin{pmatrix} (E - mc^2)I_2 & -c\vec{\sigma}\cdot\vec{p} \\ -c\vec{\sigma}\cdot\vec{p} & -(E + mc^2)I_2 \end{pmatrix}\begin{pmatrix} u_A \\ u_B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{12.706}$$

Where I_2 denotes the 2×2 unit matrix,

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \tag{12.707}$$

for notational brevity, we will be dropping in future expressions the unit matrix symbol I_2 (that was introduced in (12.706)). Now, we can rewrite (12.706) in terms of the following two coupled equations

$$(E - mc^2)u_A = c\vec{\sigma}\cdot\vec{p}u_B, \tag{12.708}$$

$$c\vec{\sigma}\cdot\vec{p}u_A = (E + mc^2)u_B. \tag{12.709}$$

These four coupled equations are nothing but an abbreviation of the Dirac equation (12.231) for a free particle as described by the plane wave (12.306).

As specified in (12.320), the free-particle equations (12.316)–(12.318) have four solutions – two linearly independent, doubly degenerate solutions with positive energy $E^{(+)}$ and two other doubly degenerate solutions corresponding to the negative energy $E^{(-)}$. In what follows, we are going to focus on determining the positive and then the negative energy solutions.

Positive Energy Solutions: Inserting (12.704) into (12.709), we obtain

$$u_B = \frac{c\vec{\sigma}\cdot\vec{p}}{E^{(+)} + mc^2}u_A = \begin{pmatrix} \frac{cp_z}{E^{(+)}+mc^2} & \frac{c(p_x - ip_y)}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} & \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix} u_A, \tag{12.710}$$

which, when combined with (12.705), leads to

$$u(p^\mu) = N\begin{pmatrix} u_A \\ u_B \end{pmatrix} = N\begin{pmatrix} u_A \\ \begin{pmatrix} \frac{cp_z}{E^{(+)}+mc^2} & \frac{c(p_x - ip_y)}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} & \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix} u_A \end{pmatrix} \tag{12.711}$$

Equation (12.710) shows that the components of u_A and u_B are interdependent. Once the components of u_A are identified, the two components of u_B are specified by (12.710), and vice versa. So, the amplitude $u(p^\mu)$ of the plane wave contains an arbitrary quantity, u_A, with two components (bi-spinor) that have yet to be determined. Since the components of u_A can be

freely chosen, we can select them by analogy to the solutions (12.296)–(12.298) of a particle at rest; that is, we can choose them to correspond to the two values of the projection of the particle's spin along, for instance, the z–axis within the particle's rest frame:

$$u_A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \text{or} \qquad u_A = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{12.712}$$

Inserting these expressions of u_A into (12.711), we obtain the two linearly independent solutions corresponding to $E^{(+)}$:

$$u_1(p^\mu) = N \begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \end{pmatrix}, \qquad u_2(p^\mu) = N \begin{pmatrix} 0 \\ 1 \\ \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix}. \tag{12.713}$$

Negative Energy Solutions: Let us now turn to the two negative energy solutions, corresponding to $E^{(-)} = -\sqrt{c^2\vec{p}^{\,2} + m^2c^4}$. Inserting (12.704) into (12.708), we have

$$u_A = \frac{c\vec{\sigma}\cdot\vec{p}}{E^{(-)} - mc^2} u_B = \begin{pmatrix} \frac{cp_z}{E^{(-)}-mc^2} & \frac{c(p_x-ip_y)}{E^{(-)}-mc^2} \\ \frac{c(p_x+ip_y)}{E^{(-)}-mc^2} & \frac{-cp_z}{E^{(-)}-mc^2} \end{pmatrix} u_B, \tag{12.714}$$

which, when combined with (12.705), leads to

$$u(p^\mu) = N \begin{pmatrix} u_A \\ u_B \end{pmatrix} = N \begin{pmatrix} \begin{pmatrix} \frac{cp_z}{E^{(-)}-mc^2} & \frac{c(p_x-ip_y)}{E^{(-)}-mc^2} \\ \frac{c(p_x+ip_y)}{E^{(-)}-mc^2} & \frac{-cp_z}{E^{(-)}-mc^2} \end{pmatrix} u_B \\ u_B \end{pmatrix}. \tag{12.715}$$

Again, we can chose the components of u_B as follows:

$$u_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \text{or} \qquad u_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{12.716}$$

Hence, inserting (12.716) into (12.715), we can infer the solutions corresponding to $E^{(-)}$

$$u_3(p^\mu) = N \begin{pmatrix} \frac{cp_z}{E^{(-)}-mc^2} \\ \frac{c(p_x+ip_y)}{E^{(-)}-mc^2} \\ 1 \\ 0 \end{pmatrix}, \qquad u_4(p^\mu) = N \begin{pmatrix} \frac{c(p_x-ip_y)}{E^{(-)}-mc^2} \\ \frac{-cp_z}{E^{(-)}-mc^2} \\ 0 \\ 1 \end{pmatrix}. \tag{12.717}$$

(b) Let us now find the factor N appearing in equations (12.713) and (12.717). Recall that, in nonrelativistic single-particle quantum mechanics, the normalization convention is quite straightforward: the probability of finding a particle somewhere in space is equal to one; i.e., the wave function, being a single-component function, is normalized to one particle per unit volume V: $\int \rho \, dV = \int \psi^* \psi \, dV = 1$. In relativistic quantum mechanics, however, the wave function is no longer a single component function, it is a spinor with four components so

as to account for the two spin orientations of the particle and its antiparticle. The standard convention is to normalize ψ to $2E$ particles per unit volume:

$$\int_{\text{unit volume}} \psi^\dagger \psi \, dV = 2E, \tag{12.718}$$

where $E > 0$. This will make the wave function Lorentz invariant, since the Lorentz contraction of the volume[81] will be compensated by a change in energy. In practice, to find N, we simply need to normalize any one of the spinors $|u_1\rangle$, $|u_2\rangle$, $|u_3\rangle$, $|u_4\rangle$; they all lead to the same value of N. Again, to find N, and to ensure adherence with causality[82], we can select a normalization condition so that the spinors $|u_1\rangle$, $|u_2\rangle$, $|u_3\rangle$, $|u_4\rangle$ transform like the time-like component of a Lorentz 4-vector with the requirement of having $2E^{(+)}$ particles per unit volume as mentioned above:

$$u_n^\dagger u_n \equiv \langle u_n | u_n \rangle = 2E^{(+)}, \qquad n = 1, 2, 3, 4. \tag{12.719}$$

Let us illustrate the calculations on $|u_1\rangle$. After taking the Hermitian adjoint of $|u_1\rangle$ from equation (12.713) and making use of the fact that p_x, p_y, and p_z are Hermitian, we have

$$\begin{aligned}
\langle u_1 | u_1 \rangle &= |N|^2 \begin{pmatrix} 1 & 0 & \frac{cp_z}{E^{(+)}+mc^2} & \frac{c(p_x - ip_y)}{E^{(+)}+mc^2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \end{pmatrix} \\
&= |N|^2 \left[1 + \frac{c^2 p_z^2}{(E^{(+)}+mc^2)^2} + \frac{c^2(p_x - ip_y)(p_x + ip_y)}{(E^{(+)}+mc^2)^2} \right] \\
&= |N|^2 \left[1 + \frac{c^2 p_z^2}{(E^{(+)}+mc^2)^2} + \frac{c^2(p_x^2 + p_y^2)}{(E^{(+)}+mc^2)^2} \right] \\
&= |N|^2 \left[1 + \frac{c^2 \vec{p}^{\,2}}{(E^{(+)}+mc^2)^2} \right] \\
&= |N|^2 \frac{(E^{(+)}+mc^2)^2 + c^2 \vec{p}^{\,2}}{(E^{(+)}+mc^2)^2}.
\end{aligned} \tag{12.720}$$

Using the relation $c^2 \vec{p}^{\,2} = (E^{(+)})^2 - m^2 c^4$, we can rewrite (12.720) as

$$\begin{aligned}
\langle u_1 | u_1 \rangle &= |N|^2 \frac{(E^{(+)}+mc^2)^2 + c^2 \vec{p}^{\,2}}{(E^{(+)}+mc^2)^2} \\
&= |N|^2 \frac{(E^{(+)})^2 + m^2 c^4 + 2mc^2 E^{(+)} + (E^{(+)})^2 - m^2 c^4}{(E^{(+)}+mc^2)^2} \\
&= |N|^2 \frac{2E^{(+)}\left(E^{(+)}+mc^2\right)}{(E^{(+)}+mc^2)^2} \\
&= |N|^2 \frac{2E^{(+)}}{E^{(+)}+mc^2},
\end{aligned} \tag{12.721}$$

[81] The volume will contract because the lengths parallel to the motion undergo Lorentz contraction.

[82] As discussed in Appendix E, time-like events can be connected by causal, real physical signals.

which, when combined with $\langle u_1|u_1\rangle = 2E^{(+)}$ as seen in (12.719), leads to

$$\boxed{N = \sqrt{E^{(+)} + mc^2}.} \tag{12.722}$$

Remarks

- We should note that there is a second normalization condition in the literature that starts from

$$\overline{u}_n u_n = u_n^\dagger \gamma^0 u_n \equiv \langle u_n|\gamma^0|u_n\rangle = 1, \qquad n = 1,\ 2,\ 3,\ 4. \tag{12.723}$$

Illustrating the calculations on u_1 and since

$$\begin{aligned}\gamma_0|u_1\rangle &= N\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}1\\0\\\frac{cp_z}{E^{(+)}+mc^2}\\\frac{c(p_x+ip_y)}{E^{(+)}+mc^2}\end{pmatrix}\\ &= N\begin{pmatrix}1\\0\\-\frac{cp_z}{E^{(+)}+mc^2}\\-\frac{c(p_x+ip_y)}{E^{(+)}+mc^2}\end{pmatrix},\end{aligned} \tag{12.724}$$

we have

$$\begin{aligned}\overline{u}_1 u_1 = \langle u_1|\gamma^0|u_1\rangle &= |N|^2\left(1 \quad 0 \quad \frac{cp_z}{E^{(+)}+mc^2} \quad \frac{c(p_x-ip_y)}{E^{(+)}+mc^2}\right)\begin{pmatrix}1\\0\\-\frac{cp_z}{E^{(+)}+mc^2}\\-\frac{c(p_x+ip_y)}{E^{(+)}+mc^2}\end{pmatrix}\\ &= |N|^2\left[1-\frac{c^2p_z^2}{(E^{(+)}+mc^2)^2}-\frac{c^2(p_x-ip_y)(p_x+ip_y)}{(E^{(+)}+mc^2)^2}\right]\\ &= |N|^2\left[1-\frac{c^2p_z^2}{(E^{(+)}+mc^2)^2}-\frac{c^2(p_x^2+p_y^2)}{(E^{(+)}+mc^2)^2}\right]\\ &= |N|^2\left[1-\frac{c^2\vec{p}^{\,2}}{(E^{(+)}+mc^2)^2}\right]\\ &= |N|^2\frac{(E^{(+)}+mc^2)^2-c^2\vec{p}^{\,2}}{(E^{(+)}+mc^2)^2}.\end{aligned} \tag{12.725}$$

Now, inserting $c^2\vec{p}^{\,2} = (E^{(+)})^2 - m^2c^4$ into (12.725), we obtain

$$\begin{aligned}1 = \overline{u}_1 u_1 &= |N|^2\frac{(E^{(+)}+mc^2)^2-c^2\vec{p}^{\,2}}{(E^{(+)}+mc^2)^2}\\ &= |N|^2\frac{(E^{(+)})^2+m^2c^4+2mc^2E^{(+)}-(E^{(+)})^2+m^2c^4}{(E^{(+)}+mc^2)^2}\\ &= |N|^2\frac{2mc^2\left(E^{(+)}+mc^2\right)}{(E^{(+)}+mc^2)^2}\\ &= |N|^2\frac{2mc^2}{E^{(+)}+mc^2},\end{aligned} \tag{12.726}$$

which leads to

$$N = \sqrt{\frac{E^{(+)} + mc^2}{2mc^2}}. \tag{12.727}$$

- Throughout this chapter, however, we use the first normalization condition (12.722): $N = \sqrt{E^{(+)} + mc^2}$.

To summarize, using equations (12.713)–(12.717), we can write the solutions of the Dirac equation for a free particle with non-zero momentum as follows:

- two doubly degenerate states with positive energy $E^{(+)} = \sqrt{c^2\vec{p}^{\,2} + m^2c^4}$

$$\psi_1(x^\mu) = N \begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \end{pmatrix} e^{-i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}, \qquad \psi_2(x^\mu) = N \begin{pmatrix} 0 \\ 1 \\ \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix} e^{-i(E^{(+)}t-\vec{p}\cdot\vec{r})/\hbar}, \tag{12.728}$$

where we have used the fact that $p_\mu x^\mu = E^{(+)}t - \vec{p}\cdot\vec{r}$.

- and two other doubly degenerate solutions with *negative energy* $E^{(-)} = -\sqrt{c^2\vec{p}^{\,2} + m^2c^4}$,

$$\psi_3(x^\mu) = N \begin{pmatrix} \frac{cp_z}{E^{(-)}-mc^2} \\ \frac{c(p_x+ip_y)}{E^{(-)}-mc^2} \\ 1 \\ 0 \end{pmatrix} e^{-i(E^{(-)}t-\vec{p}\cdot\vec{r})/\hbar}, \qquad \psi_4(x^\mu) = N \begin{pmatrix} \frac{c(p_x-ip_y)}{E^{(-)}-mc^2} \\ \frac{-cp_z}{E^{(-)}-mc^2} \\ 0 \\ 1 \end{pmatrix} e^{-i(E^{(-)}t-\vec{p}\cdot\vec{r})/\hbar}, \tag{12.729}$$

where we have used the fact that, in the case of negative energy solutions, $p_\mu x^\mu$ is given by $p_\mu x^\mu = E^{(-)}t - \vec{p}\cdot\vec{r}$.

Problem 12.13
Using the expressions of the γ matrices, find the matrix representation for a slashed quantity $\not{b} = \gamma^\mu b_\mu$ where b_μ is a 4-vector.

Solution

Inserting the matrices (12.222)–(12.223) into the following relation

$$\not{b} \equiv \gamma^\mu b_\mu = \gamma^0 b^0 - \vec{\gamma}\cdot\vec{b} = \gamma^0 b^0 - \gamma^1 b_x - \gamma^2 b_y - \gamma^3 b_z, \tag{12.730}$$

we can write

$$\begin{aligned}
\not{b} &= \gamma^0 b^0 - \gamma^1 b_x - \gamma^2 b_y - \gamma^3 b_z \\
&= b_0 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} - b_x \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \\
&\quad -b_y \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} - b_z \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} b_0 & 0 & -b_z & -b_x + ib_y \\ 0 & b_0 & -b_x - ib_y & b_z \\ b_z & b_x - ib_y & -b_0 & 0 \\ b_x + ib_y & -b_z & 0 & -b_0 \end{pmatrix}. \qquad (12.731)
\end{aligned}$$

Problem 12.14

(a) Show that the "covariant" γ–matrices, γ_0, γ_1, γ_2, γ_3, are related to their "contravariant" counterparts, γ^0, γ^1, γ^2, γ^3, by

$$\gamma_0 = \gamma^0, \qquad \gamma_1 = -\gamma^1, \qquad \gamma_2 = -\gamma^2, \qquad \gamma_3 = -\gamma^3.$$

(b) Using the above equation, write down the matrices of γ_0, γ_1, γ_2, γ_3.

Solution

(a) Let us begin by showing the first relation, $\gamma_0 = \gamma^0$. Using the relation $\gamma_\mu = g_{\mu\nu}\gamma^\nu$ (see Eq. (12.182)) along with $g_{00} = 1$, $g_{01} = 0$, $g_{02} = 0$, $g_{03} = 0$, we can easily obtain the expression of γ_0:

$$\gamma_0 = g_{0\nu}\gamma^\nu \equiv \sum_{\nu=0}^{3} g_{0\nu}\gamma^\nu = g_{00}\gamma^0 + g_{01}\gamma^1 + g_{02}\gamma^2 + g_{03}\gamma^3 = g_{00}\gamma^0 \equiv \gamma^0. \qquad (12.732)$$

Similarly, since $g_{10} = 0$, $g_{11} = -1$, $g_{12} = 0$, $g_{13} = 0$, we can easily obtain the expression of γ_1:

$$\gamma_1 = g_{1\nu}\gamma^\nu \equiv \sum_{\nu=0}^{3} g_{1\nu}\gamma^\nu = g_{10}\gamma^0 + g_{11}\gamma^1 + g_{12}\gamma^2 + g_{13}\gamma^3 = g_{11}\gamma^1 \equiv -\gamma^1. \qquad (12.733)$$

Following the same method, and since $g_{22} = -1$ and $g_{33} = -1$, we obtain at once the expressions of γ_2 and γ_3:

$$\gamma_2 = g_{2\nu}\gamma^\nu \equiv \sum_{\nu=0}^{3} g_{2\nu}\gamma^\nu = g_{20}\gamma^0 + g_{21}\gamma^1 + g_{22}\gamma^2 + g_{23}\gamma^3 = g_{22}\gamma^2 \equiv -\gamma^2, \qquad (12.734)$$

$$\gamma_3 = g_{3\nu}\gamma^\nu \equiv \sum_{\nu=0}^{3} g_{3\nu}\gamma^\nu = g_{30}\gamma^0 + g_{31}\gamma^1 + g_{32}\gamma^2 + g_{33}\gamma^3 = g_{33}\gamma^3 \equiv -\gamma^3. \qquad (12.735)$$

In summary, the "covariant" γ–matrices, γ_0, γ_1, γ_2, γ_3, are related to their "contravariant" counterparts, γ^0, γ^1, γ^2, γ^3, by the following relations:

$$\boxed{\gamma_0 = \gamma^0, \quad \gamma_1 = -\gamma^1, \quad \gamma_2 = -\gamma^2, \quad \gamma_3 = -\gamma^3.} \tag{12.736}$$

(b) Combining (12.222)–(12.223) and (12.736), we obtain the matrices of γ_0, γ_1, γ_2, γ_3:

$$\gamma_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \gamma_1 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{12.737}$$

$$\gamma_2 = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \qquad \gamma_3 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \tag{12.738}$$

Problem 12.15

(a) Calculate the inverse of the following 4×4 matrix

$$U = \begin{pmatrix} I_2 & \sigma_y \\ -\sigma_y & I_2 \end{pmatrix},$$

where I_2 is the 2×2 unity matrix and σ_y is the y–component of the Pauli matrix.

(b) Using the Majorana basis matrices (12.227) and the Dirac basis matrices (12.222) - (12.223), show that $\gamma_M^3 = U\gamma_D^3 U^{-1}$ where γ_D^3 and γ_M^3 are the third components of the Dirac matrix and the Majorana matrix, respectively.

Solution:

(a) First, a simple calculation yields the inverse of U:

$$U^{-1} = \begin{pmatrix} I_2 & \sigma_y \\ -\sigma_y & I_2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 & -i \\ 0 & 1 & i & 0 \\ 0 & i & 1 & 0 \\ -i & 0 & 0 & 1 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & i \\ 0 & 1 & -i & 0 \\ 0 & -i & 1 & 0 \\ i & 0 & 0 & 1 \end{pmatrix} \tag{12.739}$$

(b) Using the Dirac and Majorana matrices obtained from (12.227) and matrices (12.222) - (12.223),

$$\gamma_D^3 = \begin{pmatrix} 0 & \sigma_z \\ -\sigma_z & 0 \end{pmatrix}, \qquad \gamma_M^3 = \begin{pmatrix} -i\sigma_x & 0 \\ 0 & -i\sigma_x \end{pmatrix}, \tag{12.740}$$

along with (12.739), we can write

$$U\gamma_D^3 U^{-1} = \frac{1}{2}\begin{pmatrix}1&0&0&-i\\0&1&i&0\\0&i&1&0\\-i&0&0&1\end{pmatrix}\begin{pmatrix}0&0&1&0\\0&0&0&-1\\-1&0&0&0\\0&1&0&0\end{pmatrix}\begin{pmatrix}1&0&0&i\\0&1&-i&0\\0&-i&1&0\\i&0&0&1\end{pmatrix}$$
$$= \begin{pmatrix}0&-i&0&0\\-i&0&0&0\\0&0&0&-i\\0&0&-i&0\end{pmatrix}. \tag{12.741}$$

The last matrix is simply equal to the Majorana matrix γ_M^3 listed in (12.740). Hence, we can write

$$U\gamma_D^3 U^{-1} = \gamma_M^3 = \begin{pmatrix}-i\sigma_x & 0\\0 & -i\sigma_x\end{pmatrix} = \begin{pmatrix}0&-i&0&0\\-i&0&0&0\\0&0&0&-i\\0&0&-i&0\end{pmatrix}. \tag{12.742}$$

Problem 12.16
Consider the following two 4-vectors: $a^\mu = (-2, 4, -1, 2)$ and $b^\mu = (3, -2, 3, 1)$.

(a) Find the matrices representing the slashed 4-vectors: $\not{a}$ and $\not{b}$.

(b) Calculate Tr $(\not{a})$ and Tr $(\not{b})$.

(c) Calculate Tr $(\not{a}\not{b})$.

(d) Calculate Tr $(\gamma^5\not{a}\not{b})$.

Solution:

(a) Following the same method that led to Equation (12.731), using the matrices of γ^0, γ^1, γ^2, and γ^3, and using the values of $a^\mu = (-2, 4, -1, 2)$ and $b^\mu = (3, -2, 3, 1)$, we can obtain at once the matrices of $\not{a}$ and $\not{b}$:

$$\not{a} = \gamma^0 a^0 - \gamma^1 a_x - \gamma^2 a_y - \gamma^3 a_z$$
$$= \begin{pmatrix}a_0 & 0 & -a_z & -a_x + ia_y\\0 & a_0 & -a_x - ia_y & a_z\\a_z & a_x - ia_y & -a_0 & 0\\a_x + ia_y & -a_z & 0 & -a_0\end{pmatrix}$$
$$= \begin{pmatrix}-2 & 0 & -2 & -4-i\\0 & -2 & -4+i & 2\\2 & 4+i & 2 & 0\\4-i & -2 & 0 & 2\end{pmatrix}, \tag{12.743}$$

$$\not{b} = \begin{pmatrix}b_0 & 0 & -b_z & -b_x + ib_y\\0 & b_0 & -b_x - ib_y & b_z\\b_z & b_x - ib_y & -b_0 & 0\\b_x + ib_y & -b_z & 0 & -b_0\end{pmatrix} = \begin{pmatrix}3 & 0 & -1 & 2+3i\\0 & 3 & 2-3i & 1\\1 & -2-3i & -3 & 0\\-2+3i & -1 & 0 & -3\end{pmatrix}. \tag{12.744}$$

(b) From (12.743), we see that $\not{a}$ is traceless

$$\text{Tr}\,(\not{a}) = \text{Tr}\begin{pmatrix} -2 & 0 & -2 & -4-i \\ 0 & -2 & -4+i & 2 \\ 2 & 4+i & 2 & 0 \\ 4-i & -2 & 0 & 2 \end{pmatrix} = -2-2+2+2 = 0. \tag{12.745}$$

Similarly, using (12.744), we see that $\not{b}$ is also traceless

$$\text{Tr}\,(\not{b}) = \text{Tr}\begin{pmatrix} 3 & 0 & -1 & 2+3i \\ 0 & 3 & 2-3i & 1 \\ 1 & -2-3i & -3 & 0 \\ -2+3i & -1 & 0 & -3 \end{pmatrix} = 3+3-3-3 = 0. \tag{12.746}$$

This is expected since, as mentioned in Theorem (12.207), the trace of the product of an odd number of matrices is zero.

(c) Taking the product of (12.743) and (12.744)

$$\begin{aligned} \not{a}\not{b} &= \begin{pmatrix} -2 & 0 & -2 & -4-i \\ 0 & -2 & -4+i & 2 \\ 2 & 4+i & 2 & 0 \\ 4-i & -2 & 0 & 2 \end{pmatrix}\begin{pmatrix} 3 & 0 & -1 & 2+3i \\ 0 & 3 & 2-3i & 1 \\ 1 & -2-3i & -3 & 0 \\ -2+3i & -1 & 0 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 3-10i & 8+7i & 8 & 8-3i \\ -8+7i & 3+10i & 8+3i & -8 \\ 8 & 8-3i & 3-10i & 8+7i \\ 8+3i & -8 & -8+7i & 3+10i \end{pmatrix} \end{aligned} \tag{12.747}$$

Hence, the trace of $\not{a}\not{b}$ is given by:

$$\begin{aligned} \text{Tr}\,(\not{a}\not{b}) &= \text{Tr}\begin{pmatrix} 3-10i & 8+7i & 8 & 8-3i \\ -8+7i & 3+10i & 8+3i & -8 \\ 8 & 8-3i & 3-10i & 8+7i \\ 8+3i & -8 & -8+7i & 3+10i \end{pmatrix} \\ &= (3-10i)+(3+10i)+(3-10i)+(3+10i) \\ &= 12. \end{aligned} \tag{12.748}$$

This is also expected since the trace of the product of an even number of matrices is, in general, not zero.

(d) Using the matrix of γ^5 and equation (12.747), we obtain

$$\begin{aligned} \gamma^5\not{a}\not{b} &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 3-10i & 8+7i & 8 & 8-3i \\ -8+7i & 3+10i & 8+3i & -8 \\ 8 & 8-3i & 3-10i & 8+7i \\ 8+3i & -8 & -8+7i & 3+10i \end{pmatrix} \\ &= \begin{pmatrix} 8 & 8-3i & 3-10i & 8+7i \\ 8+3i & -8 & -8+7i & 3+10i \\ 3-10i & 8+7i & 8 & 8-3i \\ -8+7i & 3+10i & 8+3i & -8 \end{pmatrix} \end{aligned} \tag{12.749}$$

Hence, we have

$$\begin{aligned}\mathrm{Tr}\left(\gamma^5 \not{a}\not{b}\right) &= \mathrm{Tr}\begin{pmatrix} 8 & 8-3i & 3-10i & 8+7i \\ 8+3i & -8 & -8+7i & 3+10i \\ 3-10i & 8+7i & 8 & 8-3i \\ -8+7i & 3+10i & 8+3i & -8 \end{pmatrix} \\ &= 8-8+8-8=0. \end{aligned} \tag{12.750}$$

This is expected since the trace of the product of an odd number of matrices is zero.

Problem 12.17
Show that the S matrix, $S = \exp\left(-\frac{i}{4}\alpha_{\mu\nu}\sigma^{\mu\nu}\right)$, corresponding to a finite rotation of angle ϕ along the z–axis in the Dirac space is given by

$$S = \exp\left(-\frac{i}{2}\alpha_{12}\sigma^{12}\right) = I_4 \cos\frac{\phi}{2} - i\gamma_1\gamma_2 \sin\frac{\phi}{2}.$$

Solution:

For a rotation along the z–axis, the values of μ, ν correspond to $\mu, \nu = 1, 2$. Since, $\alpha_{\mu\nu}$ and $\sigma^{\mu\nu}$ are both antisymmetric (i.e., $\alpha_{12} = -\alpha_{21}$ and $\sigma^{12} = -\sigma^{21}$) and using the notation $\alpha_{12} = \phi$ for the finite rotation angle, we have: $\alpha_{\mu\nu}\sigma^{\mu\nu} = \alpha_{12}\sigma^{12} + \alpha_{21}\sigma^{21} = 2\alpha_{12}\sigma^{12} = 2\phi\sigma^{12}$. Hence, we can write the S matrix as follow

$$S = \exp\left(-\frac{i}{4}\alpha_{\mu\nu}\sigma^{\mu\nu}\right) = \exp\left(-\frac{i}{2}\alpha_{12}\sigma^{12}\right) = \exp\left(-\frac{i}{2}\phi\sigma^{12}\right). \tag{12.751}$$

First, since the product $\gamma^1\gamma^2$ is given by

$$\gamma^1\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix} \tag{12.752}$$

and since $\gamma^1\gamma^2 = -\gamma^2\gamma^1$, we can obtain at once the commutator $[\gamma^1, \gamma^2]$

$$[\gamma^1, \gamma^2] = \gamma^1\gamma^2 - \gamma^2\gamma^1 = 2\gamma^1\gamma^2 = 2\begin{pmatrix} -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix}. \tag{12.753}$$

Hence, we can write

$$\sigma^{12} = \frac{i}{2}\left[\gamma^1, \gamma^2\right] = i\gamma^1\gamma^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \equiv \begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}, \tag{12.754}$$

where σ_3 is the 2×2 Pauli matrix along the z–axis: $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Inserting (12.754) into (12.751), we can write S matrix as follows

$$\begin{aligned} S &= \exp\left(-\frac{i}{2}\phi\sigma^{12}\right) = \sum_{n=0}^{\infty}\frac{(-i)^{2n}}{(2n)!}\left(\frac{\phi}{2}\right)^{2n}\left(\sigma^{12}\right)^{2n} + \sum_{n=0}^{\infty}\frac{(-i)^{2n+1}}{(2n+1)!}\left(\frac{\phi}{2}\right)^{2n+1}\left(\sigma^{12}\right)^{2n+1} \\ &= \sum_{n=0}^{\infty}\frac{(-i)^{2n}}{(2n)!}\left(\frac{\phi}{2}\right)^{2n}\begin{pmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix}^{2n} + \sum_{n=0}^{\infty}\frac{(-i)^{2n+1}}{(2n+1)!}\left(\frac{\phi}{2}\right)^{2n+1}\begin{pmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix}^{2n+1}. \end{aligned} \tag{12.755}$$

Since σ_3 satisfies these properties: $\sigma_3^{2n} = I_2$ and $\sigma_3^{2n+1} = \sigma_3$ where I_2 is the 2×2 unity matrix, we can easily show that

$$\left(\sigma^{12}\right)^{2n} = \begin{pmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix}^{2n} = I_4, \qquad \left(\sigma^{12}\right)^{2n+1} = \begin{pmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix}^{2n+1} = \begin{pmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix}. \tag{12.756}$$

This can be seen as follows: using (12.754), we can write

$$\left(\sigma^{12}\right)^2 = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}^2 = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} = I_4. \tag{12.757}$$

This relations leads at once to

$$\left(\sigma^{12}\right)^{2n} = I_4^{2n} = I_4, \tag{12.758}$$

which, in turn, yields

$$\left(\sigma^{12}\right)^{2n+1} = \left(\sigma^{12}\right)^{2n}\left(\sigma^{12}\right) = I_4\sigma^{12} = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix} \equiv \begin{pmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix}. \tag{12.759}$$

Now, inserting (12.756) into (12.755), we obtain

$$S = \exp\left(-\frac{i}{2}\phi\sigma^{12}\right) = \begin{pmatrix}I_2 & 0\\ 0 & I_2\end{pmatrix}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}\left(\frac{\phi}{2}\right)^{2n} - i\begin{pmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}\left(\frac{\phi}{2}\right)^{2n+1}. \tag{12.760}$$

which reduces to

$$\boxed{S = \exp\left[-\tfrac{i}{2}\phi\begin{pmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix}\right] = \begin{pmatrix}I_2 & 0\\ 0 & I_2\end{pmatrix}\cos\left(\tfrac{\phi}{2}\right) - i\begin{pmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix}\sin\left(\tfrac{\phi}{2}\right).} \tag{12.761}$$

We can write this relation in the form of a 4×4 matrix as follows:

$$S = \exp\left(-\frac{i}{2}\phi\sigma^{12}\right) = \begin{pmatrix}\cos\left(\frac{\phi}{2}\right) - i\sin\left(\frac{\phi}{2}\right) & 0 & 0 & 0\\ 0 & \cos\left(\frac{\phi}{2}\right) + i\sin\left(\frac{\phi}{2}\right) & 0 & 0\\ 0 & 0 & \cos\left(\frac{\phi}{2}\right) - i\sin\left(\frac{\phi}{2}\right) & 0\\ 0 & 0 & 0 & \cos\left(\frac{\phi}{2}\right) + i\sin\left(\frac{\phi}{2}\right)\end{pmatrix}. \tag{12.762}$$

or as

$$S = \exp\left(-\frac{i}{2}\phi\sigma^{12}\right) = \begin{pmatrix} \exp(-i\phi/2) & 0 & 0 & 0 \\ 0 & \exp(i\phi/2) & 0 & 0 \\ 0 & 0 & \exp(-i\phi/2) & 0 \\ 0 & 0 & 0 & \exp(i\phi/2) \end{pmatrix}. \tag{12.763}$$

Problem 12.18
Show that the S matrix $S = \exp\left(-\frac{i}{4}\alpha_{\mu\nu}\sigma^{\mu\nu}\right)$ corresponding to a finite boost χ along the $x-$axis in the Dirac space is given by

$$S = \exp\left(-\frac{i}{2}\alpha_{01}\sigma^{01}\right) = I_4 \cosh\frac{\chi}{2} + \gamma_0\gamma_1 \sinh\frac{\chi}{2}.$$

Solution:

For a boost along the $x-$axis, the values of μ, ν correspond to $\mu, \nu = 0, 1$. Since, $\alpha_{\mu\nu}$ and $\sigma^{\mu\nu}$ are both antisymmetric (i.e., $\alpha_{01} = -\alpha_{10}$ and $\sigma^{01} = -\sigma^{10}$) and using the notation $\alpha_{01} = \chi$ for the finite boost parameter, we have: $\alpha_{\mu\nu}\sigma^{\mu\nu} = \alpha_{01}\sigma^{01} + \alpha_{10}\sigma^{10} = 2\alpha_{01}\sigma^{01} = 2\chi\sigma^{01}$. Hence, we can write the S matrix as follow

$$S = \exp\left(-\frac{i}{4}\alpha_{\mu\nu}\sigma^{\mu\nu}\right) = \exp\left(-\frac{i}{2}\alpha_{01}\sigma^{01}\right) = \exp\left(-\frac{i}{2}\chi\sigma^{01}\right). \tag{12.764}$$

First, since the product $\gamma^0\gamma^1$ is given by

$$\gamma^0\gamma^1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{12.765}$$

and since $\gamma^1\gamma^0 = -\gamma^0\gamma^1$, we can obtain at once the commutator $[\gamma^0, \gamma^1]$:

$$[\gamma^0, \gamma^1] = \gamma^0\gamma^1 - \gamma^1\gamma^0 = 2\gamma^0\gamma^1 = 2\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{12.766}$$

as well as the expression of σ^{01}:

$$\sigma^{01} = \frac{i}{2}[\gamma^0, \gamma^1] = i\gamma^0\gamma^1 = i\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \equiv i\begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}, \tag{12.767}$$

where σ_1 is the 2×2 Pauli matrix along the $x-$axis: $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Inserting (12.767) into (12.764), we can write S matrix as follows

$$\begin{aligned} S &= \exp\left(-\frac{i}{2}\chi\sigma^{01}\right) = \exp\left[\frac{1}{2}\chi\begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}\right] \\ &= \sum_{n=0}^{\infty}\frac{1}{(2n)!}\left(\frac{\chi}{2}\right)^{2n}\begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}^{2n} + \sum_{n=0}^{\infty}\frac{1}{(2n+1)!}\left(\frac{\chi}{2}\right)^{2n+1}\begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}^{2n+1}. \end{aligned} \tag{12.768}$$

Now, since σ_1 satisfy these properties: $\sigma_1^{2n} = I_2$ and $\sigma_1^{2n+1} = \sigma_1$ where I_2 is the 2×2 unity matrix, we can easily show that

$$\left(\sigma^{01}\right)^{2n} = \begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}^{2n} = I_4, \qquad \left(\sigma^{01}\right)^{2n+1} = \begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}^{2n+1} = \begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}. \tag{12.769}$$

The proof of these two relations goes as follows. Using (12.767), we can write

$$\begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}^2 = \begin{pmatrix}0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\end{pmatrix}^2 = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} = I_4. \tag{12.770}$$

This relations leads at once to

$$\begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}^{2n} = I_4^{2n} = I_4, \tag{12.771}$$

which, in turn, yields

$$\begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}^{2n+1} = \begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}^{2n}\begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix} = \begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}. \tag{12.772}$$

Now, inserting (12.769) into (12.768), we obtain

$$S = \exp\left[\frac{1}{2}\chi\begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}\right] = \begin{pmatrix}I_2 & 0\\ 0 & I_2\end{pmatrix}\sum_{n=0}^{\infty}\frac{(1}{(2n)!}\left(\frac{\chi}{2}\right)^{2n} + \begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}\left(\frac{\chi}{2}\right)^{2n+1}. \tag{12.773}$$

which reduces to

$$\boxed{S = \exp\left[\tfrac{1}{2}\chi\begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}\right] = \begin{pmatrix}I_2 & 0\\ 0 & I_2\end{pmatrix}\cosh\left(\tfrac{\chi}{2}\right) + \begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}\sinh\left(\tfrac{\chi}{2}\right).} \tag{12.774}$$

We can write this relation in the form of a 4×4 matrix as follows:

$$S = \exp\left[\frac{1}{2}\chi\begin{pmatrix}0 & \sigma_1\\ \sigma_1 & 0\end{pmatrix}\right] = \begin{pmatrix}\cosh\left(\frac{\chi}{2}\right) & 0 & 0 & \sinh\left(\frac{\chi}{2}\right)\\ 0 & \cosh\left(\frac{\chi}{2}\right) & \sinh\left(\frac{\chi}{2}\right) & 0\\ 0 & \sinh\left(\frac{\chi}{2}\right) & \cosh\left(\frac{\chi}{2}\right) & 0\\ \sinh\left(\frac{\chi}{2}\right) & 0 & 0 & \cosh\left(\frac{\chi}{2}\right)\end{pmatrix}. \tag{12.775}$$

Problem 12.19
Consider a matrix quantity, Σ, which is defined in terms of Dirac's γ matrices as follows

$$\Sigma^{\mu\nu} = \frac{i}{2}[\gamma^{\mu}, \gamma^{\nu}].$$

(a) Using the matrix representations of the γ matrices, find the matrices representing

$$\Sigma_x = \Sigma^{23}, \qquad \Sigma_y = \Sigma^{31}, \qquad \Sigma_z = \Sigma^{12}.$$

(b) Show that Σ_x, Σ_y, and Σ_z satisfy the same commutation relations as the 2×2 Pauli matrices.

(c) Show that the matrix of $\vec{\Sigma} = \Sigma_x\hat{i} + \Sigma_y\hat{j} + \Sigma_z\hat{k}$ reduces to

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix},$$

where $\vec{\sigma} = \sigma_x\hat{i} + \sigma_y\hat{j} + \sigma_z\hat{k}$ are the 2×2 Pauli matrices listed in (12.162).

(d) Find the matrix of the helicity operator, $\widehat{\mathcal{H}} = (\hbar/2p)\vec{\Sigma}\cdot\vec{p}$ where $\vec{p}$ is the linear momentum.

Solution

(a) Using the γ–matrices listed in Eqs. (12.222) — (12.223), we can obtain

$$\begin{aligned}
\Sigma_x &= \frac{i}{2}[\gamma^2, \gamma^3] = \frac{i}{2}(\gamma^2\gamma^3 - \gamma^3\gamma^2) \\
&= \frac{i}{2}\left(\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}\right) \\
&= \frac{i}{2}\left(\begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \end{pmatrix} - \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & i & 0 \end{pmatrix}\right) = \frac{i}{2}\begin{pmatrix} 0 & -2i & 0 & 0 \\ -2i & 0 & 0 & 0 \\ 0 & 0 & 0 & -2i \\ 0 & 0 & -2i & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},
\end{aligned} \tag{12.776}$$

$$\begin{aligned}
\Sigma_y &= \frac{i}{2}\left[\gamma^3, \gamma^1\right] = \frac{i}{2}\left(\gamma^3\gamma^1 - \gamma^3\gamma^1\right) \\
&= \frac{i}{2}\left(\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\right) \\
&= \frac{i}{2}\left(\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}\right) = \frac{i}{2}\begin{pmatrix} 0 & -2 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 2 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix},
\end{aligned} \tag{12.777}$$

$$\begin{aligned}
\Sigma_z &= \frac{i}{2}\left[\gamma^1, \gamma^2\right] = \frac{i}{2}\left(\gamma^1\gamma^2 - \gamma^2\gamma^1\right) \\
&= \frac{i}{2}\left(\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}\right) \\
&= \frac{i}{2}\left(\begin{pmatrix} -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix} - \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{pmatrix}\right) = \frac{i}{2}\begin{pmatrix} -2i & 0 & 0 & 0 \\ 0 & 2i & 0 & 0 \\ 0 & 0 & -2i & 0 \\ 0 & 0 & 0 & 2i \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.
\end{aligned} \tag{12.778}$$

(b) Let us now calculate the three commutation relations between Σ_x, Σ_y, Σ_z derived in Eqs. (12.776) — (12.778):

$$\begin{aligned}
\left[\Sigma_x, \Sigma_y\right] &= \Sigma_x\Sigma_y - \Sigma_y\Sigma_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\
&= 2i\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.
\end{aligned} \tag{12.779}$$

Comparing (12.776) and (12.779), we see that

$$\left[\Sigma_x, \Sigma_y\right] = 2i\Sigma_z. \tag{12.780}$$

Following the method above, we can easily obtain the other two commutators

$$\begin{aligned}\left[\Sigma_y,\ \Sigma_z\right] &= \Sigma_y\Sigma_z - \Sigma_z\Sigma_y = \begin{pmatrix}0 & -i & 0 & 0\\ i & 0 & 0 & 0\\ 0 & 0 & 0 & -i\\ 0 & 0 & i & 0\end{pmatrix}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix} - \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}\begin{pmatrix}0 & -i & 0 & 0\\ i & 0 & 0 & 0\\ 0 & 0 & 0 & -i\\ 0 & 0 & i & 0\end{pmatrix}\\ &\equiv 2i\begin{pmatrix}0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{pmatrix}\\ &\equiv 2i\Sigma_x \qquad (12.781)\end{aligned}$$

$$\begin{aligned}[\Sigma_z,\ \Sigma_x] &= \Sigma_z\Sigma_x - \Sigma_x\Sigma_z = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}\begin{pmatrix}0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{pmatrix} - \begin{pmatrix}0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{pmatrix}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}\\ &= \begin{pmatrix}0 & 2 & 0 & 0\\ -2 & 0 & 0 & 0\\ 0 & 0 & 0 & 2\\ 0 & 0 & -2 & 0\end{pmatrix} = 2i\begin{pmatrix}0 & -i & 0 & 0\\ i & 0 & 0 & 0\\ 0 & 0 & 0 & -i\\ 0 & 0 & i & 0\end{pmatrix}\\ &\equiv 2i\Sigma_y \qquad (12.782)\end{aligned}$$

So, a combination of (12.780) — (12.782) shows that Σ_x, Σ_y, and Σ_z satisfy the same commutation relations as the 2×2 Pauli matrices; namely:

$$\left[\Sigma_x,\ \Sigma_y\right] = 2i\Sigma_z, \qquad \left[\Sigma_y,\ \Sigma_z\right] = 2i\Sigma_x, \qquad [\Sigma_z,\ \Sigma_x] = 2i\Sigma_y. \qquad (12.783)$$

(c) Using Eqs. (12.776) — (12.778), we can easily express the $\vec{\Sigma} = \Sigma_x\hat{i} + \Sigma_y\hat{j} + \Sigma_z\hat{k}$ in the following matrix form:

$$\begin{aligned}\vec{\Sigma} &= \Sigma_x\hat{i} + \Sigma_y\hat{j} + \Sigma_z\hat{k}\\ &= \begin{pmatrix}0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{pmatrix}\hat{i} + \begin{pmatrix}0 & -i & 0 & 0\\ i & 0 & 0 & 0\\ 0 & 0 & 0 & -i\\ 0 & 0 & i & 0\end{pmatrix}\hat{j} + \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}\hat{k}\\ &= \begin{pmatrix}\hat{k} & \hat{i}-i\hat{j} & 0 & 0\\ \hat{i}+i\hat{j} & -\hat{k} & 0 & 0\\ 0 & 0 & \hat{k} & \hat{i}-i\hat{j}\\ 0 & 0 & \hat{i}+i\hat{j} & -\hat{k}\end{pmatrix}. \qquad (12.784)\end{aligned}$$

A combination of (12.162) and (12.784) yields:

$$\vec{\Sigma} = \begin{pmatrix}\hat{k} & \hat{i}-i\hat{j} & 0 & 0\\ \hat{i}+i\hat{j} & -\hat{k} & 0 & 0\\ 0 & 0 & \hat{k} & \hat{i}-i\hat{j}\\ 0 & 0 & \hat{i}+i\hat{j} & -\hat{k}\end{pmatrix} \equiv \begin{pmatrix}\vec{\sigma} & 0\\ 0 & \vec{\sigma}\end{pmatrix}, \qquad (12.785)$$

since

$$\vec{\sigma} = \sigma_x \hat{i} + \sigma_y \hat{j} + \sigma_z \hat{k} = \begin{pmatrix} \hat{k} & \hat{i} - i\hat{j} \\ \hat{i} + i\hat{j} & -\hat{k} \end{pmatrix}. \tag{12.786}$$

(d) Using (12.785), we can obtain at once the matrix of the helicity operator $(\hbar/2p)\vec{\Sigma} \cdot \vec{p}$:

$$\begin{aligned} \widehat{\mathcal{H}} &= \frac{\hbar}{2p}\vec{\Sigma} \cdot \vec{p} = \frac{\hbar}{2p} \begin{pmatrix} \hat{k} & \hat{i} - i\hat{j} & 0 & 0 \\ \hat{i} + i\hat{j} & -\hat{k} & 0 & 0 \\ 0 & 0 & \hat{k} & \hat{i} - i\hat{j} \\ 0 & 0 & \hat{i} + i\hat{j} & -\hat{k} \end{pmatrix} \cdot \vec{p} \\ &= \frac{\hbar}{2p} \begin{pmatrix} p_z & p_x - ip_y & 0 & 0 \\ p_x + ip_y & -p_z & 0 & 0 \\ 0 & 0 & p_z & p_x - ip_z \\ 0 & 0 & p_x + itp_y & -p_z \end{pmatrix} \\ &\equiv \frac{\hbar}{2p} \begin{pmatrix} \vec{\sigma} \cdot \vec{p} & 0 \\ 0 & \vec{\sigma} \cdot \vec{p} \end{pmatrix} \end{aligned} \tag{12.787}$$

where $p = \sqrt{p_x^2 + p_y^2 + p_z^2}$ is the magnitude of the momentum.

Problem 12.20
Consider the states $|\psi_1(t)\rangle$, $|\psi_2(t)\rangle$, $|\psi_3(t)\rangle$, $|\psi_4(t)\rangle$ listed in equations (12.297)–(12.298).

(a) Verify that the states $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$ correspond to spins up and down with values $+\hbar/2$ and $-\hbar/2$, respectively.

(b) Repeat (a) for the states $|\psi_3(t)\rangle$ and $|\psi_4(t)\rangle$; namely, verify that they correspond to spins up and down with values $+\hbar/2$ and $-\hbar/2$, respectively.

Solution

(a) Since we are dealing with a free particle in a 4×4 space, the spin operator $\hat{S}_z$ is given by

$$\hat{S}_z = \frac{\hbar}{2} \begin{pmatrix} \sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{12.788}$$

where σ_z is the z–component of the Pauli matrices

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{12.789}$$

First, applying $\hat{S}_z$ on $|\psi_1(t)\rangle$ listed in equation (12.297), we have

$$\begin{aligned}\hat{S}_z|\psi_1(t)\rangle &= \hat{S}_z\begin{pmatrix}1\\0\\0\\0\end{pmatrix}e^{-imc^2t/\hbar} = \frac{\hbar}{2}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}e^{-imc^2t/\hbar}\\ &= \frac{\hbar}{2}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}e^{-imc^2t/\hbar}\\ &\equiv \frac{\hbar}{2}|\psi_1(t)\rangle. \end{aligned} \tag{12.790}$$

Hence, being an eigenstate of $\hat{S}_z$, the state $|\psi_1(t)\rangle$ clearly describes a spin-up particle with an $\frac{\hbar}{2}$ value. Similarly, applying $\hat{S}_z$ on $|\psi_2(t)\rangle$ listed in equation (12.297), we have

$$\begin{aligned}\hat{S}_z|\psi_2(t)\rangle &= \hat{S}_z|\begin{pmatrix}0\\1\\0\\0\end{pmatrix}e^{-imc^2t/\hbar} = \frac{\hbar}{2}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}e^{-imc^2t/\hbar}\\ &= -\frac{\hbar}{2}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}e^{-imc^2t/\hbar}\\ &\equiv -\frac{\hbar}{2}|\psi_2(t)\rangle. \end{aligned} \tag{12.791}$$

Hence, as $|\psi_2(t)\rangle$ is an eigenstate of $\hat{S}_z$, it does describe a spin-down particle with a $-\frac{\hbar}{2}$ value.

(b) second, applying $\hat{S}_z$ on $|\psi_3(t)\rangle$ listed in equation (12.298), we can easily verify that

$$\hat{S}_z|\psi_3(t)\rangle = \frac{\hbar}{2}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}e^{imc^2t/\hbar} = \frac{\hbar}{2}|\psi_3(t)\rangle; \tag{12.792}$$

hence, $|\psi_3(t)\rangle$ corresponds to a spin-up with an $\frac{\hbar}{2}$ value. Similarly, applying $\hat{S}_z$ on $|\psi_4(t)\rangle$ listed in equation (12.298), we have

$$\hat{S}_z|\psi_4(t)\rangle = -\frac{\hbar}{2}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}e^{imc^2t/\hbar} = -\frac{\hbar}{2}|\psi_4(t)\rangle, \tag{12.793}$$

which shows that $|\psi_4(t)\rangle$ corresponds to a spin-down with a $-\frac{\hbar}{2}$ value.

Problem 12.21

Show that the Dirac Hamiltonian $\hat{H} = c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2$ commutes with the helicity operator $\widehat{\mathcal{H}} = \hat{\vec{S}}\cdot\left(\frac{\vec{p}}{|\vec{p}|}\right)$ where $\hat{\vec{S}} = \frac{\hbar}{2}\begin{pmatrix}\vec{\sigma}&0\\0&\vec{\sigma}\end{pmatrix}$ where $\sigma_x, \sigma_y, \sigma_z$ are the 2×2 Pauli matrices.

Solution:

First, let us derive the matrix of $\widehat{\mathcal{H}}$. Inserting the 2×2 Pauli matrices (12.162) into $\widehat{\mathcal{H}}$, we obtain

$$\begin{aligned}\widehat{\mathcal{H}} &= \frac{\hbar}{2}\begin{pmatrix}\vec{\sigma} & 0\\ 0 & \vec{\sigma}\end{pmatrix}\cdot\left(\frac{\vec{p}}{|\vec{p}|}\right) = \frac{\hbar}{2}\left(\frac{p_x}{|\vec{p}|}\right)\begin{pmatrix}\sigma_x & 0\\ 0 & \sigma_x\end{pmatrix} + \frac{\hbar}{2}\left(\frac{p_y}{|\vec{p}|}\right)\begin{pmatrix}\sigma_y & 0\\ 0 & \sigma_y\end{pmatrix} + \frac{\hbar}{2}\left(\frac{p_z}{|\vec{p}|}\right)\begin{pmatrix}\sigma_z & 0\\ 0 & \sigma_x\end{pmatrix}\\ &\equiv \frac{\hbar}{2}\frac{p_x}{|\vec{p}|}\begin{pmatrix}0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{pmatrix} + \frac{\hbar}{2}\frac{p_y}{|\vec{p}|}\begin{pmatrix}0 & -i & 0 & 0\\ i & 0 & 0 & 0\\ 0 & 0 & 0 & -i\\ 0 & 0 & i & 0\end{pmatrix} + \frac{\hbar}{2}\frac{p_z}{|\vec{p}|}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix},\end{aligned} \tag{12.794}$$

or

$$\boxed{\widehat{\mathcal{H}} = \frac{\hbar}{2|\vec{p}|}\begin{pmatrix}p_z & p_x - ip_y & 0 & 0\\ p_x + ip_y & -p_z & 0 & 0\\ 0 & 0 & p_z & p_x - ip_z\\ 0 & 0 & p_x + ip_y & -p_z\end{pmatrix}.} \tag{12.795}$$

Second, let us obtain the matrix of the Dirac Hamiltonian. Using the matrices of β, α_x, α_y, α_z listed in equations (12.161) and (12.163), we can easily obtain the matrix of the Dirac Hamiltonian

$$\begin{aligned}\hat{H} &= c\vec{\alpha}\cdot\hat{\vec{p}} + mc^2\beta = c(\alpha_x p_x + \alpha_y p_y + \alpha_z p_z) + \beta mc^2\\ &= \begin{pmatrix}mc^2 & 0 & cp_z & c(p_x - ip_y)\\ 0 & mc^2 & c(p_x + ip_y) & -cp_z\\ cp_z & c(p_x - ip_y) & -mc^2 & 0\\ c(p_x + ip_y) & -cp_z & 0 & -mc^2\end{pmatrix}.\end{aligned} \tag{12.796}$$

Third, combining equations (12.795) and (12.796), we can obtain the product of the Hamiltonian $\hat{H}$ with the helicity $\widehat{\mathcal{H}}$:

$$\begin{aligned}\hat{H}\widehat{\mathcal{H}} &= \frac{\hbar}{2|\vec{p}|}\begin{pmatrix}mc^2 & 0 & cp_z & c(p_x - ip_y)\\ 0 & mc^2 & c(p_x + ip_y) & -cp_z\\ cp_z & c(p_x - ip_y) & -mc^2 & 0\\ c(p_x + ip_y) & -cp_z & 0 & -mc^2\end{pmatrix}\begin{pmatrix}p_z & p_x - ip_y & 0 & 0\\ p_x + ip_y & -p_z & 0 & 0\\ 0 & 0 & p_z & p_x - ip_z\\ 0 & 0 & p_x + ip_y & -p_z\end{pmatrix}\\ &= \frac{\hbar}{2|\vec{p}|}\begin{pmatrix}mc^2p_z & mc^2(p_x - ip_y) & cp_z^2 + c(p_x^2 + p_y^2) & 0\\ mc^2(p_x + ip_y) & -mc^2p_z & 0 & cp_z^2 + c(p_x^2 + p_y^2)\\ cp_z^2 + c(p_x^2 + p_y^2) & 0 & -mc^2p_z & -mc^2(p_x - ip_y)\\ 0 & cp_z^2 + c(p_x^2 + p_y^2) & -mc^2(p_x + ip_y) & mc^2p_z\end{pmatrix}\end{aligned} \tag{12.797}$$

Similarly, the product of the helicity $\widehat{\mathcal{H}}$ with the Hamiltonian $\hat{H}$ is given by:

$$\widehat{\mathcal{H}}\,\hat{H} = \frac{\hbar}{2|\vec{p}|}\begin{pmatrix} p_z & p_x - ip_y & 0 & 0 \\ p_x + ip_y & -p_z & 0 & 0 \\ 0 & 0 & p_z & p_x - ip_z \\ 0 & 0 & p_x + ip_y & -p_z \end{pmatrix}\begin{pmatrix} mc^2 & 0 & cp_z & c(p_x - ip_y) \\ 0 & mc^2 & c(p_x + ip_y) & -cp_z \\ cp_z & c(p_x - ip_y) & -mc^2 & 0 \\ c(p_x + ip_y) & -cp_z & 0 & -mc^2 \end{pmatrix}$$

$$= \frac{\hbar}{2|\vec{p}|}\begin{pmatrix} mc^2 p_z & mc^2(p_x - ip_y) & cp_z^2 + c(p_x^2 + p_y^2) & 0 \\ mc^2(p_x + ip_y) & -mc^2 p_z & 0 & cp_z^2 + c(p_x^2 + p_y^2) \\ cp_z^2 + c(p_x^2 + p_y^2) & 0 & -mc^2 p_z & -mc^2(p_x - ip_y) \\ 0 & cp_z^2 + c(p_x^2 + p_y^2) & -mc^2(p_x + ip_y) & mc^2 p_z \end{pmatrix} \tag{12.798}$$

Finally, comparing equations (12.797) and (12.798), we see that the Dirac Hamiltonian commutes with helicity operator: $\hat{H}\widehat{\mathcal{H}} = \widehat{\mathcal{H}}\hat{H}$; that is,

$$\boxed{[\hat{H}, \widehat{\mathcal{H}}] = \left[c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2,\ \hat{\vec{S}}\cdot\left(\frac{\vec{p}}{|\vec{p}|}\right)\right] = 0.} \tag{12.799}$$

Problem 12.22
Using the matrix expression (12.795), show that the eigenvalues of the helicity operator $\widehat{\mathcal{H}}$ are $\pm 1/2$.

Solution:
Using the matrix of $\widehat{\mathcal{H}}$ listed in Eq. (12.795), we can write:

$$\widehat{\mathcal{H}}^2 = \frac{\hbar^2}{4|\vec{p}|^2}\begin{pmatrix} p_z & p_x - ip_y & 0 & 0 \\ p_x + ip_y & -p_z & 0 & 0 \\ 0 & 0 & p_z & p_x - ip_z \\ 0 & 0 & p_x + ip_y & -p_z \end{pmatrix}\begin{pmatrix} p_z & p_x - ip_y & 0 & 0 \\ p_x + ip_y & -p_z & 0 & 0 \\ 0 & 0 & p_z & p_x - ip_z \\ 0 & 0 & p_x + ip_y & -p_z \end{pmatrix}$$

$$= \frac{\hbar^2}{4|\vec{p}|^2}\begin{pmatrix} (p_x^2 + p_y^2 + p_z^2) & 0 & 0 & 0 \\ 0 & (p_x^2 + p_y^2 + p_z^2) & 0 & 0 \\ 0 & 0 & (p_x^2 + p_y^2 + p_z^2) & 0 \\ 0 & 0 & 0 & (p_x^2 + p_y^2 + p_z^2) \end{pmatrix}$$

$$= \frac{\hbar^2}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \frac{\hbar^2}{4} I_4. \tag{12.800}$$

This shows that the eigenvalues of $\widehat{\mathcal{H}}$ are $\pm\hbar/2$.

Problem 12.23
Show that the Dirac Hamiltonian $\hat{H} = c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2$ does not commute with the operator $\hat{\vec{\Sigma}} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}$ where $\sigma_x, \sigma_y, \sigma_z$ are the 2×2 Pauli matrices.

Solution:

The commutator between $\hat{H}$ and $\hat{\vec{\Sigma}}$ is given by

$$[\hat{H},\ \hat{\vec{\Sigma}}] = [c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2,\ \hat{\vec{\Sigma}}] = c[\vec{\alpha}\cdot\hat{\vec{p}},\ \hat{\vec{\Sigma}}] + mc^2[\beta,\ \hat{\vec{\Sigma}}] \equiv c[\vec{\alpha}\cdot\hat{\vec{p}},\ \hat{\vec{\Sigma}}]. \tag{12.801}$$

In the last step we have used the fact that $[\beta,\ \hat{\vec{\Sigma}}] = 0$ since, as displayed in (12.161), the matrix of β is diagonal. Now, let us begin by calculating the commutators $[\hat{H},\ \hat{\Sigma}_x]$ and then calculate the commutators corresponding to $\hat{\Sigma}_y$ and $\hat{\Sigma}_z$. Since the momentum operator $\hat{\vec{p}}$ does not act on $\vec{\alpha}$ nor on $\hat{\vec{\Sigma}}$, we can write

$$\begin{aligned}[\hat{H},\ \hat{\Sigma}_x] &= c[\alpha_x\beta_x + \alpha_y\beta_y + \alpha_z\beta_z,\ \hat{\Sigma}_x] \\ &= c\beta_x[\alpha_x,\ \hat{\Sigma}_x] + c\beta_y[\alpha_y,\ \hat{\Sigma}_x] + c\beta_z[\alpha_z,\ \hat{\Sigma}_x].\end{aligned} \tag{12.802}$$

Using the 2×2 representations of $\vec{\alpha}$ and $\hat{\vec{\Sigma}}$

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \qquad \vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}, \tag{12.803}$$

we can can easily verify that $[\alpha_x,\ \hat{\Sigma}_x] = 0$:

$$\begin{aligned}[\alpha_x,\ \hat{\Sigma}_x] &= \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix}\begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} - \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix}\begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \sigma_x^2 \\ \sigma_x^2 & 0 \end{pmatrix} - \begin{pmatrix} 0 & \sigma_x^2 \\ \sigma_x^2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ &= 0.\end{aligned} \tag{12.804}$$

Following the same method, and using the well known commutation relation between the Pauli matrices $[\sigma_x,\ \sigma_y] = 2i\sigma_z$, we can write

$$\begin{aligned}[\alpha_y,\ \hat{\Sigma}_x] &= \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix}\begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} - \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix}\begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \sigma_y\sigma_x - \sigma_x\sigma_y \\ \sigma_y\sigma_x - \sigma_x\sigma_y & 0 \end{pmatrix} \equiv \begin{pmatrix} 0 & -[\sigma_x,\ \sigma_y] \\ -[\sigma_x,\ \sigma_y] & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -2i\sigma_z \\ -2i\sigma_z & 0 \end{pmatrix} = -2i\begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} \\ &= -2i\alpha_z.\end{aligned} \tag{12.805}$$

Similarly, using the commutator $[\sigma_z,\ \sigma_x] = 2i\sigma_y$, we can obtain

$$\begin{aligned}[\alpha_z,\ \hat{\Sigma}_x] &= \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix}\begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} - \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix}\begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \sigma_z\sigma_x - \sigma_x\sigma_z \\ \sigma_z\sigma_x - \sigma_x\sigma_z & 0 \end{pmatrix} = \begin{pmatrix} 0 & [\sigma_z,\ \sigma_x] \\ [\sigma_z,\ \sigma_x] & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 2i\sigma_y \\ 2i\sigma_y & 0 \end{pmatrix} \\ &= 2i\alpha_y.\end{aligned} \tag{12.806}$$

Substituting equations (12.804), (12.805), (12.806) into (12.802), we obtain

$$[\hat{H},\ \hat{\Sigma}_x] = c\hat{p}_x\left[\alpha_x,\ \hat{\Sigma}_x\right] + c\hat{p}_y\left[\alpha_y,\ \hat{\Sigma}_x\right] + c\hat{p}_z\left[\alpha_z,\ \hat{\Sigma}_x\right] = 2ic(\hat{p}_z\alpha_y - \hat{p}_y\alpha_z). \tag{12.807}$$

Since the term $(\hat{p}_z\alpha_y - \hat{p}_y\alpha_z)$ is simply the $x-$component of the cross product between $\vec{\alpha}$ and $\hat{\vec{p}}$, we can write

$$[\hat{H},\ \hat{\Sigma}_x] = 2ic(\hat{p}_z\alpha_y - \hat{p}_y\alpha_z) \equiv 2ic(\ \vec{\alpha} \times \hat{\vec{p}}\ )_x. \tag{12.808}$$

Next, following the same method outlined above, we can easily obtain the commutator $[\hat{H},\ \hat{\Sigma}_y]$. That is:

$$[\hat{H},\ \hat{\Sigma}_y] = c\hat{p}_x[\alpha_x,\ \hat{\Sigma}_y] + c\hat{p}_y[\alpha_y,\ \hat{\Sigma}_y] + c\hat{p}_z[\alpha_z,\ \hat{\Sigma}_y]. \tag{12.809}$$

We can can easily verify that $[\alpha_y,\ \hat{\Sigma}_y] = 0$:

$$\left[\alpha_y,\ \hat{\Sigma}_y\right] = \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix}\begin{pmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{pmatrix} - \begin{pmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{pmatrix}\begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix} = 0.$$

Now, using the commutation relation $[\sigma_x,\ \sigma_y] = 2i\sigma_z$, we can write

$$\begin{aligned}
[\alpha_x,\ \hat{\Sigma}_y] &= \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix}\begin{pmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{pmatrix} - \begin{pmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{pmatrix}\begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & \sigma_x\sigma_y - \sigma_y\sigma_x \\ \sigma_x\sigma_y - \sigma_y\sigma_x & 0 \end{pmatrix} \equiv \begin{pmatrix} 0 & [\sigma_x,\ \sigma_y] \\ [\sigma_x,\ \sigma_y] & 0 \end{pmatrix} \\
&= 2i\begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} \\
&= 2i\alpha_z.
\end{aligned} \tag{12.810}$$

Similarly, using the commutator $[\sigma_z,\ \sigma_y] = -2i\sigma_x$, we can obtain

$$\begin{aligned}
[\alpha_z,\ \hat{\Sigma}_y] &= \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix}\begin{pmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{pmatrix} - \begin{pmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{pmatrix}\begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & \sigma_z\sigma_y - \sigma_y\sigma_z \\ \sigma_z\sigma_y - \sigma_y\sigma_z & 0 \end{pmatrix} = \begin{pmatrix} 0 & [\sigma_z,\ \sigma_y] \\ [\sigma_z,\ \sigma_y] & 0 \end{pmatrix} \\
&= -2i\begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix} \\
&= -2i\alpha_x.
\end{aligned} \tag{12.811}$$

Substituting equations (12.810), (12.810), (12.811) into (12.809), we obtain

$$[\hat{H},\ \hat{\Sigma}_y] = c\hat{p}_x[\alpha_x,\ \hat{\Sigma}_y] + c\hat{p}_y[\alpha_y,\ \hat{\Sigma}_y] + c\hat{p}_z[\alpha_z,\ \hat{\Sigma}_y] = 2ic(\hat{p}_x\alpha_z - \hat{p}_z\alpha_x). \tag{12.812}$$

Again, since the term $(\hat{p}_x\alpha_z - \hat{p}_z\alpha_x)$ is simply the $y-$component of the cross product between $\vec{\alpha}$ and $\hat{\vec{p}}$, we can write

$$[\hat{H},\ \hat{\Sigma}_y] = 2ic(\hat{p}_x\alpha_z - \hat{p}_z\alpha_x) \equiv 2ic(\ \vec{\alpha} \times \hat{\vec{p}}\ )_y. \tag{12.813}$$

Following the same method, we can show that

$$[\hat{H},\ \hat{\Sigma}_z] = 2ic(\hat{p}_y\alpha_x - \hat{p}_x\alpha_y) \equiv 2ic(\ \vec{\alpha} \times \hat{\vec{p}}\ )_z. \tag{12.814}$$

Finally, combining equations (12.808), (12.813) and (12.814), we see that the Dirac Hamiltonian $\hat{H} = c\vec{\alpha}\cdot\hat{\vec{p}} + \beta mc^2$ does not commute with the operator $\hat{\vec{\Sigma}}$:

$$\boxed{[\hat{H},\ \hat{\vec{\Sigma}}] = 2ic(\vec{\alpha}\times\hat{\vec{p}}).} \tag{12.815}$$

Problem 12.24
Using the matrices of S_x, S_y, and S_z listed in (12.384), calculate
(a) the commutation relations between S_x, S_y, and S_z, and
(b) the matrix of $\vec{S}^2$.

Solution:

(a) Using the matrices (12.384), obtain

$$\begin{aligned}
S_xS_y - S_yS_x &= \frac{\hbar^2}{4}\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}\begin{pmatrix}0&-i&0&0\\i&0&0&0\\0&0&0&-i\\0&0&i&0\end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix}0&-i&0&0\\i&0&0&0\\0&0&0&-i\\0&0&i&0\end{pmatrix}\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix} \\
&= \frac{\hbar^2}{4}\begin{pmatrix}i&0&0&0\\0&-i&0&0\\0&0&i&0\\0&0&0&-i\end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix}-i&0&0&0\\0&i&0&0\\0&0&-i&0\\0&0&0&i\end{pmatrix} = \frac{\hbar^2}{4}\begin{pmatrix}2i&0&0&0\\0&-2i&0&0\\0&0&2i&0\\0&0&0&-2i\end{pmatrix} \\
&= i\hbar\frac{\hbar}{2}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix} \\
&= i\hbar S_z.
\end{aligned} \tag{12.816}$$

Next, a simple calculation leads to

$$\begin{aligned}
S_yS_z - S_zS_y &= \frac{\hbar^2}{4}\begin{pmatrix}0&-i&0&0\\i&0&0&0\\0&0&0&-i\\0&0&i&0\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0&-i&0&0\\i&0&0&0\\0&0&0&-i\\0&0&i&0\end{pmatrix} \\
&= \frac{\hbar^2}{4}\begin{pmatrix}0&i&0&0\\i&0&0&0\\0&0&0&i\\0&0&i&0\end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix}0&-i&0&0\\-i&0&0&0\\0&0&0&-i\\0&0&-i&0\end{pmatrix} = \frac{\hbar^2}{4}\begin{pmatrix}0&2i&0&0\\2i&0&0&0\\0&0&0&2i\\0&0&2i&0\end{pmatrix} \\
&= i\hbar\frac{\hbar}{2}\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix} \\
&= i\hbar S_x.
\end{aligned} \tag{12.817}$$

Similarly, we can easily show that

$$
\begin{aligned}
S_zS_x - S_xS_z &= \frac{\hbar^2}{4}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix} \\
&= \frac{\hbar^2}{4}\begin{pmatrix}0&1&0&0\\-1&0&0&0\\0&0&0&1\\0&0&-1&0\end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix}0&-1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&1&0\end{pmatrix} = \frac{\hbar^2}{4}\begin{pmatrix}0&2&0&0\\-2&0&0&0\\0&0&0&2\\0&0&-2&0\end{pmatrix} \\
&= \frac{\hbar^2}{4}\begin{pmatrix}0&-2i^2&0&0\\2i^2&0&0&0\\0&0&0&-2i^2\\0&0&2i^2&0\end{pmatrix} = i\hbar\frac{\hbar}{2}\begin{pmatrix}0&-i&0&0\\i&0&0&0\\0&0&0&-i\\0&0&i&0\end{pmatrix} \\
&= i\hbar S_y. \qquad (12.818)
\end{aligned}
$$

Combining equations (12.816), (12.817), and (12.818), we can obtain the following commutation relations

$$
\boxed{[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z, \qquad [\hat{S}_y, \hat{S}_z] = i\hbar\hat{S}_x, \qquad [\hat{S}_z, \hat{S}_x] = i\hbar\hat{S}_y,} \qquad (12.819)
$$

(b) Using the matrices (12.384), we obtain

$$
\begin{aligned}
\vec{S}^2 &= S_x^2 + S_y^2 + S_z^2 \\
&= \frac{\hbar^2}{4}\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}^2 + \frac{\hbar^2}{4}\begin{pmatrix}0&-i&0&0\\i&0&0&0\\0&0&0&-i\\0&0&i&0\end{pmatrix}^2 + \frac{\hbar^2}{4}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}^2 \\
&= \frac{\hbar^2}{4}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix} + \frac{\hbar^2}{4}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix} + \frac{\hbar^2}{4}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix} + \\
&= \frac{\hbar^2}{4}\begin{pmatrix}3&0&0&0\\0&3&0&0\\0&0&3&0\\0&0&0&3\end{pmatrix} = \frac{3}{4}\hbar^2\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix} \\
&\equiv \frac{3}{4}\hbar^2 I_4 \qquad (12.820)
\end{aligned}
$$

where I_4 is the 4×4 unit matrix.

Problem 12.25
Using equations (12.321), (12.334), (12.335) and (12.383), show that the Dirac spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$, in general, are not eigenstates of the operator $\hat{S}_z$ listed in equation (12.383).

Solution

Using the expression (12.321) of $|u_1\rangle$ and the matrix of $\hat{S}_z$ listed in the text (see (12.383)), we have

$$\hat{S}_z|u_1\rangle = N\frac{\hbar}{2}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}1\\0\\\frac{cp_z}{E^{(+)}+mc^2}\\\frac{c(p_x+ip_y)}{E^{(+)}+mc^2}\end{pmatrix} = N\frac{\hbar}{2}\begin{pmatrix}1\\0\\\frac{cp_z}{E^{(+)}+mc^2}\\-\frac{c(p_x+ip_y)}{E^{(+)}+mc^2}\end{pmatrix} \neq \frac{\hbar}{2}|u_1\rangle, \tag{12.821}$$

where $E^{(+)} = \sqrt{c^2\vec{p}^{\,2} + m^2c^4}$; this equation clearly shows that $|u_1\rangle$ is *not* an eigenstate of $\hat{S}_z$.

Following the same method and using equations (12.321) and (12.383), we can verify that $|u_2\rangle$ is not an eigenstate of $\hat{S}_z$

$$\hat{S}_z|u_2\rangle = N\frac{\hbar}{2}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0\\1\\\frac{c(p_x-ip_y)}{E^{(+)}+mc^2}\\\frac{-cp_z}{E^{(+)}+mc^2}\end{pmatrix} = N\frac{\hbar}{2}\begin{pmatrix}0\\-1\\\frac{c(p_x-ip_y)}{E^{(+)}+mc^2}\\\frac{cp_z}{E^{(+)}+mc^2}\end{pmatrix} \neq \frac{\hbar}{2}|u_2\rangle. \tag{12.822}$$

Now, let us look at the action of $\hat{S}_z$ on the *antiparticle* spinors v_1 and v_2. First, it is important to note that the sign of $\hat{S}_z$ changes sign when we deal with antiparticle spinors:

$$\hat{S}_z^{(A)} = -\hat{S}_z. \tag{12.823}$$

Hence

$$\hat{S}_z^{(A)}|v_1\rangle = -\hat{S}_z|v_1\rangle = -N\frac{\hbar}{2}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}\frac{c(p_x-ip_y)}{E^{(+)}+mc^2}\\\frac{-cp_z}{E^{(+)}+mc^2}\\0\\1\end{pmatrix} = -N\frac{\hbar}{2}\begin{pmatrix}\frac{c(p_x-ip_y)}{E^{(+)}+mc^2}\\\frac{cp_z}{E^{(+)}+mc^2}\\0\\-1\end{pmatrix} \neq \frac{\hbar}{2}|v_1\rangle. \tag{12.824}$$

Clearly, $|v_1\rangle$ is not an eigenstate of $\hat{S}_z^{(A)}$. Similarly, we can verify that $|v_1\rangle$ is not an eigenstate of $\hat{S}_z^{(A)}$

$$\hat{S}_z^{(A)}|v_2\rangle = -\hat{S}_z|v_2\rangle = -N\frac{\hbar}{2}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}\frac{cp_z}{E^{(+)}+mc^2}\\\frac{c(p_x+ip_y)}{E^{(+)}+mc^2}\\1\\0\end{pmatrix} = -N\frac{\hbar}{2}\begin{pmatrix}\frac{cp_z}{E^{(+)}+mc^2}\\-\frac{c(p_x+ip_y)}{E^{(+)}+mc^2}\\-1\\0\end{pmatrix} \neq \frac{\hbar}{2}|v_2\rangle. \tag{12.825}$$

Problem 12.26

Show that the Dirac spinors for a particle at rest are eigenstates of the parity operator and find the correspnding eigenvalues.

Solution:

Inserting the values $p_x = p_y = p_z = 0$ (particle at rest) into Eq. (12.336), we obtain the corresponding Dirac spinors

$$|u_1\rangle = N\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \qquad |u_2\rangle = N\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \qquad |v_1\rangle = N\begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \qquad |v_2\rangle = N\begin{pmatrix}0\\0\\1\\0\end{pmatrix}. \tag{12.826}$$

A straightforward application of the parity operator (12.547) on these spinors leads to

$$\hat{\mathcal{P}}|u_1\rangle = \pm N\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \pm N\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \pm|u_1\rangle. \tag{12.827}$$

Following the same procedure, we obtain

$$\hat{\mathcal{P}}|u_2\rangle = \pm N\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \pm|u_2\rangle, \tag{12.828}$$

and

$$\hat{\mathcal{P}}|v_1\rangle = \pm N\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \mp|v_1\rangle, \qquad \hat{\mathcal{P}}|v_2\rangle = \pm N\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \mp|v_2\rangle, \tag{12.829}$$

Hence, the Dirac spinors for a particle at rest are eigenvalues are eigenstates of the parity operator:

$$\boxed{\hat{\mathcal{P}}|u_1\rangle = \pm|u_1\rangle, \qquad \hat{\mathcal{P}}|u_2\rangle = \pm|u_2\rangle, \qquad \hat{\mathcal{P}}|v_1\rangle = \mp|v_1\rangle, \qquad \hat{\mathcal{P}}|v_2\rangle = \mp|v_2\rangle,} \tag{12.830}$$

where $|u_1\rangle$ and $|u_2\rangle$ are particle spinors, while $|v_1\rangle$ and $|v_2\rangle$ are antiparticle spinors. This result shows that *a particle at rest has an opposite intrinsic parity to an antipartilcle at rest.*

Problem 12.27
Find how the Dirac spinors corresponding to a free particle at rest transform under charge conjugation.

Solution:

First, to obtain the Dirac spinors corresponding to a free particle at rest, we would need to insert $p_x = p_y = p_z = 0$ into equations (12.337) and (12.338); the spinors corresponding to the

spin-up and spin-down particle are given by

$$\psi_1(x^\mu) = N\begin{pmatrix}1\\0\\0\\0\end{pmatrix}e^{-imc^2t/\hbar}, \qquad \psi_2(x^\mu) = N\begin{pmatrix}0\\1\\0\\0\end{pmatrix}e^{-imc^2t/\hbar}, \tag{12.831}$$

while the spin-up and spin-down antiparticle spinors are:

$$\Phi_3(x^\mu) = N\begin{pmatrix}0\\0\\0\\1\end{pmatrix}e^{imc^2t/\hbar}, \qquad \Phi_4(x^\mu) = N\begin{pmatrix}0\\0\\1\\0\end{pmatrix}e^{imc^2t/\hbar}, \tag{12.832}$$

where N is the normalization constant.

Applying the charge conjugation operator (see (12.502)),

$$\hat{C} = \begin{pmatrix}0 & 0 & 0 & 1\\ 0 & 0 & -1 & 0\\ 0 & -1 & 0 & 0\\ 1 & 0 & 0 & 0\end{pmatrix}\hat{K}, \tag{12.833}$$

to $\psi_1(x^\mu)$, we obtain

$$\begin{aligned}\hat{C}\psi_1(x^\mu) &= \begin{pmatrix}0 & 0 & 0 & 1\\ 0 & 0 & -1 & 0\\ 0 & -1 & 0 & 0\\ 1 & 0 & 0 & 0\end{pmatrix}\hat{K}\psi_1(x^\mu) = \begin{pmatrix}0 & 0 & 0 & 1\\ 0 & 0 & -1 & 0\\ 0 & -1 & 0 & 0\\ 1 & 0 & 0 & 0\end{pmatrix}\psi_1^*(x^\mu)\\ &= N\begin{pmatrix}0 & 0 & 0 & 1\\ 0 & 0 & -1 & 0\\ 0 & -1 & 0 & 0\\ 1 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}e^{imc^2t/\hbar}\\ &= N\begin{pmatrix}0\\0\\0\\1\end{pmatrix}e^{imc^2t/\hbar}\\ &\equiv \Phi_3(x^\mu).\end{aligned} \tag{12.834}$$

This shows that under the charge conjugation operator, the particle spinor $\psi_1(x^\mu)$ transforms into the antiparticles spinor $\Phi_3(x^\mu)$; for instance, the charge conjugation of a spin-up electron (charge $-e$) yields a spin-up positron (charge $+e$).

Similarly, applying $\hat{C}$ on $\psi_2(x^\mu)$ of (12.831), we obtain

$$\begin{aligned}
\hat{C}\psi_2(x^\mu) &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \hat{K}\psi_2(x^\mu) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \hat{K}\psi_2^*(x^\mu) \\
&= N \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} e^{imc^2 t/\hbar} \\
&= -N \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{imc^2 t/\hbar} \\
&\equiv -\Phi_4(x^\mu), \qquad (12.835)
\end{aligned}$$

This shows that a spin-down electron transforms under charge conjugation into a spin-down positron.

In summary, *charge conjugation transforms each particle into its antiparticle and vice versa.*

Problem 12.28
Consider a free spin-1/2 particle which is moving along the z–axis with a momentum $\vec{p} = (0, 0, p)$. Show that its Hamiltonian is invariant under time reversal.

Solution:

First, the Dirac Hamiltonian for a such a particle can be obtained by simply inserting $p_x = p_y = 0$ and $p_z = p$ in Eq. (12.314):

$$\begin{aligned}
\hat{H} &= \left. \begin{pmatrix} E - mc^2 & 0 & -cp_z & -c(p_x - ip_y) \\ 0 & E - mc^2 & -c(p_x + ip_y) & cp_z \\ cp_z & c(p_x - ip_y) & -(E + mc^2) & 0 \\ c(p_x + ip_y) & -cp_z & 0 & -(E + mc^2) \end{pmatrix} \right|_{p_x = p_y = 0,\ p_z = p} \\
&= \begin{pmatrix} E - mc^2 & 0 & -cp & 0 \\ 0 & E - mc^2 & 0 & cp \\ cp & 0 & -(E + mc^2) & 0 \\ 0 & -cp & 0 & -(E + mc^2) \end{pmatrix} \qquad (12.836)
\end{aligned}$$

Second, using the matrix (12.581) of the time reversal operator $\hat{T}$, we can calculate $\hat{T}^*\hat{H}\hat{T}$ as

follows

$$\begin{aligned}\hat{T}\hat{H}\hat{T} &= \begin{pmatrix}0 & -i & 0 & 0\\ i & 0 & 0 & 0\\ 0 & 0 & 0 & -i\\ 0 & 0 & i & 0\end{pmatrix}\begin{pmatrix}E-mc^2 & 0 & -cp & 0\\ 0 & E-mc^2 & 0 & cp\\ cp & 0 & -(E+mc^2) & 0\\ 0 & -cp & 0 & -(E+mc^2)\end{pmatrix}\begin{pmatrix}0 & i & 0 & 0\\ -i & 0 & 0 & 0\\ 0 & 0 & 0 & i\\ 0 & 0 & -i & 0\end{pmatrix}\\ &= \begin{pmatrix}E-mc^2 & 0 & -cp & 0\\ 0 & E-mc^2 & 0 & cp\\ cp & 0 & -(E+mc^2) & 0\\ 0 & -cp & 0 & -(E+mc^2)\end{pmatrix}\\ &= \hat{H}. \end{aligned} \tag{12.837}$$

Hence, *the Dirac Hamiltonian is invariant under time reversal.*

Problem 12.29

Consider a spin-up (relativistic) free electron with an energy E traveling along the positive $z-$axis and strikes a step potential of height V_0,

$$V(z) = \begin{cases} 0, & z < 0, \\ V_0, & z \geq 0, \end{cases} \tag{12.838}$$

such that $V_0 < E - mc^2$ and $-mc^2 < V_0 < mc^2$, where mc^2 is the rest-mass energy of the electron, as illustrated in Figure 12.7.

(a) Find the solutions of the Dirac equation for the electron; obtain the relations between the normalization constants of the incident, reflected, and transmitted wave functions.

(b) Find the incident, reflected, and transmitted probability densities.

(c) Find the incident, reflected, and transmitted current densities. Is the current density conserved?

(d) Find the reflection and transmission coefficients. Do they add up to 1?

Solution

(a) Since we are dealing with a one-dimensional problem where electrons are moving along the positive $z-$axis, we can rewrite the Dirac equation (12.311) as follows

$$\left(i\hbar\gamma^0\partial_0 + i\hbar\gamma^3\partial_3 - mc\right)\Psi(z,\ t) = 0; \tag{12.839}$$

Since the potential is time-independent, the time and spatial components of the solutions can be separated:

$$\Psi(z,\ t) = e^{-iEt/\hbar}\psi(z). \tag{12.840}$$

Inserting $\Psi(z,\ t)$ into (12.839) and cancelling the time-dependent exponents, $e^{-iEt/\hbar}$, we obtain two equations which involve only the time-independent wave function $\psi(z)$:

$$\left(\gamma^0\frac{E}{c} - mc\right)\psi(z) + i\hbar\gamma^3\frac{d\psi(z)}{dz} = 0, \qquad (z < 0), \tag{12.841}$$

$$\left(\gamma^0\frac{E - V_0}{c} - mc\right)\psi(z) + i\hbar\gamma^3\frac{d\psi(z)}{dz} = 0, \qquad (z \geq 0). \tag{12.842}$$

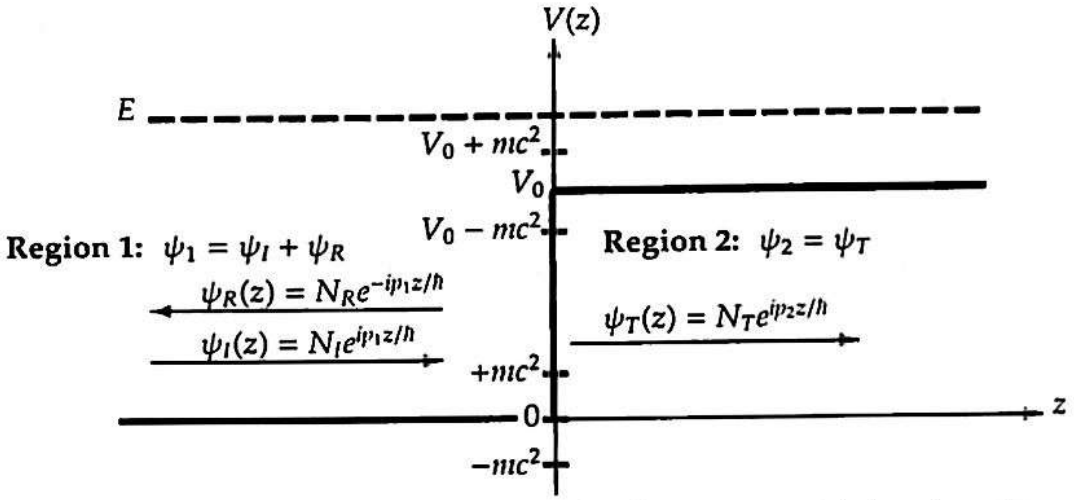


Figure 12.7: The incident, reflected, and transmitted waves pertaining to a free electron moving to the right with an energy $E > V_0 + mc^2$ and striking a step potential V_0.

As illustrated in Figure 12.7, while the solution of (12.841) consists of incident and reflected waves,

$$\psi_1(z) = \psi_I(z) + \psi_R(z), \qquad (z < 0), \tag{12.843}$$

the solution of (12.842) consists of a transmitted wave propagating to the right (no wave gets reflected from the $z > 0$ region to the left):

$$\psi_2(z) = \psi_T(z), \qquad (z \geq 0). \tag{12.844}$$

For simplicity, we have considered only one orientation of the electron's spin — spin up. To find the incident spin-up solutions, we need simply to insert $p_x = p_y = 0$ and[83] $p_z \equiv p$ in the first relation of equation (12.728); this leads to

$$\psi_I(z) = N_I \begin{pmatrix} 1 \\ 0 \\ \frac{cp_1}{E+mc^2} \\ 0 \end{pmatrix} e^{ip_1 z/\hbar}, \tag{12.845}$$

where N_I is a normalization constant for the incident wave function, and where $E = \sqrt{p_1^2 c^2 + m^2 c^4}$ and $p_1 = \sqrt{E^2/c^2 - m^2 c^2}$; p_1 is the z–component of the particle's momentum in Region 1. As for the reflected solution, it must be spin-up since no spin flip can occur from a scalar electrostatic potential (i.e, V_0); the spin is conserved at the barrier. We can obtain the reflected spin-up solution from (12.845) by simply reversing the sign of the momentum p_1:

$$\psi_R(z) = N_R \begin{pmatrix} 1 \\ 0 \\ \frac{-cp_1}{E+mc^2} \\ 0 \end{pmatrix} e^{-ip_1 z/\hbar}. \tag{12.846}$$

[83] For notational brevity throughout this problem, we will be denoting the momentum p_z by p, since the momentum has one component only.

When the electron enters into the $z > 0$ region, its momentum changes (due to the presence of the constant potential V_0); we will be denoting the electron momentum by p_2 in Region 2. So, when we scatter a free electron off a step potential V_0 located at $z = 0$, the respective 4-vector momenta corresponding to the incident, reflected, and transmitted waves are given by

$$p_I^\mu = \left(\frac{E}{c}, 0, 0, p_1\right), \qquad p_R^\mu = \left(\frac{E}{c}, 0, 0, -p_1\right), \qquad p_T^\mu = \left(\frac{E_2}{c}, 0, 0, p_2\right), \tag{12.847}$$

where E_2 and p_2 are the particle's energy and momentum in Region 2, respectively; they are given by:

$$E_2 = E - V_0, \qquad p_2 = \sqrt{\frac{E_2^2}{c^2} - m^2c^2} = \sqrt{\frac{(E - V_0)^2}{c^2} - m^2c^2}; \tag{12.848}$$

We now can obtain the transmitted spin-up solution by merely replacing E and p_1 by E_2 and p_2 in equation (12.845)

$$\psi_T(z) = N_T \begin{pmatrix} 1 \\ 0 \\ \frac{cp_2}{E_2+mc^2} \\ 0 \end{pmatrix} e^{ip_2z/\hbar}. \tag{12.849}$$

Since p_2 is real, the transmitted wave function has an oscillatory pattern; this oscillating solution for the Dirac equation in Region 2 was expected for our case of an incident electron with an energy $E > V_0 + mc^2$. If, however, we had an incident electron with an energy $V_0 - mc^2 < E < V_0 + mc^2$, the momentum p' would become imaginary and, as seen in Subsection 12.4.12, the transmitted wave function $\psi_T(z)$ decays away exponentially in Region 2.

Combining (12.843), (12.844), (12.845), (12.846), and (12.849), we can write the time-independent parts of the wave function as follows:

$$\psi(z) = \begin{cases} \psi_1(z) = N_I \begin{pmatrix} 1 \\ 0 \\ \frac{cp}{E+mc^2} \\ 0 \end{pmatrix} e^{ip_1z/\hbar} + N_R \begin{pmatrix} 1 \\ 0 \\ \frac{-cp_1}{E+mc^2} \\ 0 \end{pmatrix} e^{-ip_2z/\hbar} & z < 0 \\ \psi_2(z) = N_T \begin{pmatrix} 1 \\ 0 \\ \frac{cp_2}{E_2+mc^2} \\ 0 \end{pmatrix} e^{ip_2z/\hbar} & z \geq 0. \end{cases} \tag{12.850}$$

So, the incident and reflected waves are plane waves traveling to the right and left, respectively, in Region 1, whereas the transmitted wave is a plane wave traveling to the right in Region 2.

Now that we have found the wave functions in regions 1 and 2, we can obtain the relation between the constants N_I, N_R, and N_T from the boundary conditions. Since the Dirac equation is first order in the time derivative, we would need only one boundary condition; namely, the continuity of the Dirac's 4-spinnor at $z = 0$:

$$\psi_1(0) = \psi_2(0) \qquad \longrightarrow \qquad \psi_I(0) + \psi_R(0) = \psi_T(0) \tag{12.851}$$

Combining (12.850) with (12.851), we obtain the following matrix relation

$$N_I\begin{pmatrix}1\\0\\\frac{cp_1}{E+mc^2}\\0\end{pmatrix}+N_R\begin{pmatrix}1\\0\\\frac{-cp_1}{E+mc^2}\\0\end{pmatrix}=N_T\begin{pmatrix}1\\0\\\frac{cp_2}{E_2+mc^2}\\0\end{pmatrix}. \tag{12.852}$$

Now, by equating the top and third components of this matrix equation, respectively, we obtain the following two conditions that must be satisfied jointly:

$$N_I+N_R=N_T, \tag{12.853}$$

$$(N_I-N_R)\frac{cp_1}{E+mc^2}=N_T\frac{cp_2}{E_2+mc^2} \quad\Longrightarrow\quad N_I-N_R=\lambda N_T, \tag{12.854}$$

where the dimensionless constant λ is given by

$$\lambda=\frac{p_2}{p_1}\frac{(E+mc^2)}{(E_2+mc^2)}=\frac{p_2}{p_1}\frac{(E+mc^2)}{(E-V_0+mc^2)}=\sqrt{\frac{(E+mc^2)(E-V_0-mc^2)}{(E-mc^2)(E-V_0+mc^2)}}. \tag{12.855}$$

Multiplying equations (12.853) and (12.854), we have

$$N_I^2-N_R^2=\lambda N_T^2. \tag{12.856}$$

After dividing equations (12.853) and (12.854) by N_I, respectively, we can rewrite them as follows

$$1+\frac{N_R}{N_I}=\frac{N_T}{N_I}, \tag{12.857}$$

$$1-\frac{N_R}{N_I}=\lambda\frac{N_T}{N_I}, \tag{12.858}$$

Adding equations (12.857) and (12.858), we obtain

$$2=(1+\lambda)\frac{N_T}{N_I} \quad\Longrightarrow\quad \frac{N_T}{N_I}=\frac{2}{1+\lambda}. \tag{12.859}$$

Combining (12.857) and (12.859), we obtain

$$\frac{N_R}{N_I}=\frac{N_T}{N_I}-1=\frac{2}{1+\lambda}-1 \quad\Longrightarrow\quad \frac{N_R}{N_I}=\frac{1-\lambda}{1+\lambda}. \tag{12.860}$$

(b) Since, as shown in equation (12.280), the probability density is given by $\rho=\psi^\dagger\psi$, we can easily obtain the incident, reflected, and transmitted probability densities, respectively, by simply inserting (12.845), (12.846) and (12.849) into (12.280):

$$\rho_I = \psi_I^\dagger\psi_I=|N_I|^2\begin{pmatrix}1 & 0 & \frac{cp_1}{E+mc^2} & 0\end{pmatrix}\begin{pmatrix}1\\0\\\frac{cp_1}{E+mc^2}\\0\end{pmatrix}=\left[1+\frac{c^2p_1^2}{(E+mc^2)^2}\right]|N_I|^2$$

$$= \frac{2E}{E+mc^2}|N_I|^2, \tag{12.861}$$

$$\rho_R = \psi_R^\dagger \psi_R = |N_R|^2 \begin{pmatrix} 1 & 0 & \frac{-cp_1}{E+mc^2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{-cp_1}{E+mc^2} \\ 0 \end{pmatrix} = \left[1 + \frac{c^2 p_1^2}{(E+mc^2)^2} \right] |N_R|^2$$

$$= \frac{2E}{E+mc^2} |N_R|^2, \tag{12.862}$$

$$\rho_T = \psi_T^\dagger \psi_T = |N_T|^2 \begin{pmatrix} 1 & 0 & \frac{cp_2}{E-V_0+mc^2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{cp_2}{E-V_0+mc^2} \\ 0 \end{pmatrix} = \left[1 + \frac{c^2 p_2^2}{(E-V_0+mc^2)^2} \right] |N_T|^2$$

$$= \frac{2(E-V_0)}{E-V_0+mc^2} |N_T|^2. \tag{12.863}$$

We should note that, as seen in Subsection 12.4.12, when the incident electron has an energy in the range $V_0 - mc^2 < E < V_0 + mc^2$, the momentum p_2 becomes imaginary and, hence, the transmitted probability becomes negative. This case is associated with the creation of a particle-antiparticle pair as seen in Subsection 12.4.12.

(c) Since, as shown in equation (12.280), the probability current density is given by $\vec{j} = c\psi^\dagger \vec{\alpha} \psi$ and since the electron is moving along the $z-$axis, we would need only to calculate the $z-$component of $\vec{j}$; namely, $j_z = c\psi^\dagger \alpha_z \psi$. For notational brevity, we will be dropping the z subscript from j_z and use j instead. Hence, inserting the matrix of α_z, which is listed in (12.163), into $j = c\psi^\dagger \alpha_z \psi$ and using (12.845), we can easily obtain the incident current density

$$j_I = c\psi_I^\dagger \alpha_z \psi_I = c|N_I|^2 \begin{pmatrix} 1 & 0 & \frac{cp_1}{E+mc^2} & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{cp_1}{E+mc^2} \\ 0 \end{pmatrix}$$

$$= \frac{2c^2 p_1}{E+mc^2} |N_I|^2. \tag{12.864}$$

Similarly, inserting the matrix of α_z into $j = c\psi^\dagger \alpha_z \psi$ and using (12.846) and (12.849), respectively, we obtain

$$j_R = c\psi_R^\dagger \alpha_z \psi_R = c|N_R|^2 \begin{pmatrix} 1 & 0 & \frac{-cp_1}{E+mc^2} & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{-cp_1}{E+mc^2} \\ 0 \end{pmatrix}$$

$$= \frac{-2c^2 p_1}{E+mc^2} |N_R|^2, \tag{12.865}$$

and

$$j_T = c\psi_T^\dagger \alpha_z \psi_T = c|N_T|^2 \begin{pmatrix} 1 & 0 & \frac{cp_2}{E-V_0+mc^2} & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{cp_2}{E-V_0+mc^2} \\ 0 \end{pmatrix}$$

$$= \frac{2c^2 p_2}{E - V_0 + mc^2} |N_T|^2. \tag{12.866}$$

It helps to note that p_2 can be complex depending on the strength of V_0 (i.e., the height of the potential); as seen in Subsection 12.4.12, when the incident electron has an energy $V_0 - mc^2 < E < V_0 + mc^2$, the momentum p_2 becomes imaginary and, hence, the transmitted current density j_T also becomes complex.

Combining equations (12.864)–(12.866), we can easily obtain the following two relations

$$\frac{j_R}{j_I} = -\left|\frac{N_R}{N_I}\right|^2, \tag{12.867}$$

$$\frac{j_T}{j_I} = \frac{2p_2(E + mc^2)}{2p(E - V_0 + mc^2)} \left|\frac{N_T}{N_I}\right|^2 = \lambda \left|\frac{N_T}{N_I}\right|^2. \tag{12.868}$$

Now, substituting (12.860) into (12.867) and (12.859) into (12.868), we obtain

$$\frac{j_R}{j_I} = -\left(\frac{1-\lambda}{1+\lambda}\right)^2, \tag{12.869}$$

$$\frac{j_T}{j_I} = \frac{4\lambda}{(1+\lambda)^2}. \tag{12.870}$$

These two relations show that the current density is conserved:

$$j_I + j_R = j_T. \tag{12.871}$$

To see this, we only need to combine (12.869) and (12.870) as follows

$$1 + \frac{j_R}{j_I} = 1 - \left(\frac{1-\lambda}{1+\lambda}\right)^2 = \frac{4\lambda}{(1+\lambda)^2} \equiv \frac{j_T}{j_I} \qquad \Longrightarrow \qquad j_I + j_R = j_T. \tag{12.872}$$

(d) The *reflection* and *transmission coefficients*, R and T, are defined by

$$R = \left|\frac{\text{reflected current density}}{\text{incident current density}}\right| = \left|\frac{j_R}{j_I}\right|, \qquad T = \left|\frac{\text{transmitted current density}}{\text{incident current density}}\right| = \left|\frac{j_T}{j_I}\right|. \tag{12.873}$$

Inserting (12.869) and (12.870) into (12.873), we can easily obtain the R and T coefficients:

$$R = \left|\frac{j_R}{j_I}\right| = \left(\frac{1-\lambda}{1+\lambda}\right)^2, \qquad T = \left|\frac{j_T}{j_I}\right| = \frac{4\lambda}{(1+\lambda)^2}. \tag{12.874}$$

Since $E > V_0 + mc^2$, and using (12.855), we can easily verify that λ is real and larger than 1, $\lambda > 1$; hence, the coefficients R and T are both smaller than 1. Now, adding the two relations (12.874), we see that R and T add up to 1:

$$R + T = \left(\frac{1-\lambda}{1+\lambda}\right)^2 + \frac{4\lambda}{(1+\lambda)^2} = \frac{(1+\lambda)^2}{(1+\lambda)^2} = 1. \tag{12.875}$$

In summary, when $E > V_0 + mc^2$, we obtain the same results as in the non-relativistic quantum mechanics case; namely, some of the beam or wavepacket is reflected and some transmitted. We also have $|j_R| < |j_I|$ as expected.

Problem 12.30

Consider a spinless (relativistic) particle with an energy E traveling along the positive $z-$axis and striking a barrier potential of height V_0,

$$V(z) = \begin{cases} 0, & z \le 0, \\ V_0, & 0 < z < a, \\ 0, & z \ge a, \end{cases} \tag{12.876}$$

such that $-mc^2 < V_0 < mc^2$, where mc^2 is the rest-mass energy of the particle.

Discuss the solutions of the Klein–Gordon equation for the following energy ranges:

(a) $E > V_0 + mc^2$,

(b) $E < V_0 - mc^2$.

Solution

(a) As illustrated in left-side of Figure 12.8, we need to solve the Klein–Gordon equation in three regions: (a) Region 1, corresponding to $z \le 0$, (b) Region 2, corresponding to $0 < z < a$, and (c) Region 3, corresponding to $z > a$. Since we have a one-dimensional problem where the incident particles is moving along the positive $z-$axis, we can rewrite the Klein–Gordon equation (12.50) as follows

$$\left(\frac{d^2}{dz^2} - \frac{m^2c^2}{\hbar^2}\right)\Psi(z,\, t) = \frac{1}{c^2}\frac{\partial^2\Psi(z,\, t)}{\partial t^2}. \tag{12.877}$$

While the particle is free in regions 1 and 3, it will be moving under the influence of the constant potential V_0 in Region 2. Since the barrier potential is time-independent, the time and spatial components of the solutions in all three regions are expected to separate:

$$\Psi(z,\, t) = e^{-iEt/\hbar}\psi(z). \tag{12.878}$$

Inserting $\Psi(z,\, t)$ into (12.877) and cancelling the time-dependent exponent, $e^{-iEt/\hbar}$, we obtain two equations that correspond to regions 1 and 2, respectively, which involve only the time-independent wave function $\psi(z)$:

$$\left(\frac{d^2}{dz^2} - \frac{m^2c^2}{\hbar^2}\right)\psi(z) = -\frac{E^2}{\hbar^2c^2}\psi(z), \qquad (z \le 0 \text{ and } z \ge a), \tag{12.879}$$

$$\left(\frac{d^2}{dz^2} - \frac{m^2c^2}{\hbar^2}\right)\psi(z) = -\frac{(E-V_0)^2}{\hbar^2c^2}\psi(z), \qquad (0 < z < a). \tag{12.880}$$

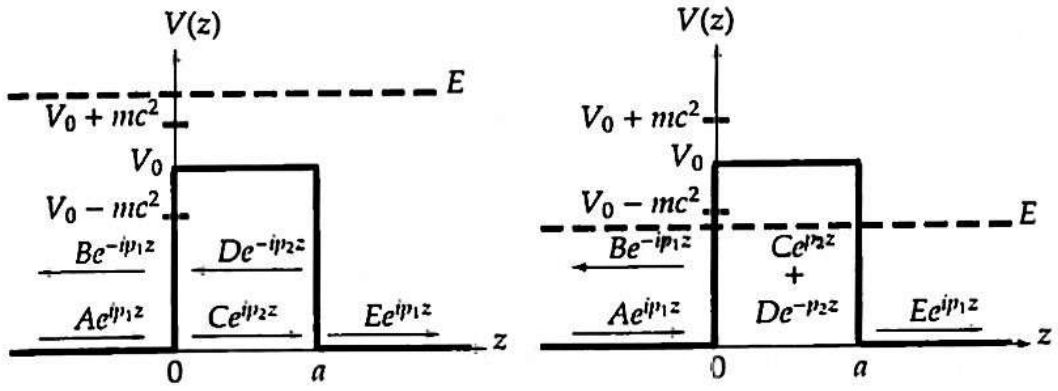


Figure 12.8: Potential barrier and propagation directions of the incident, reflected, and transmitted waves when $E > V_0 + mc^2$ and $E < V_0 - mc^2$.

In the case where $E > V_0 + mc^2$, the solutions in all three regions are planes waves:

$$\psi(z) = \begin{cases} \psi_1(z) = Ae^{ip_1z/\hbar} + Be^{-ip_1z/\hbar}, & z \le 0, \\ \psi_2(z) = Ce^{ip_2z/\hbar} + De^{-ip_2z/\hbar}, & 0 < z < a, \\ \psi_3(z) = Ee^{ip_1z/\hbar}, & a \ge a, \end{cases} \tag{12.881}$$

where the momenta in Regions 1 and 3 are equal and positive, $p_1 = p_3 = \sqrt{E^2/c^2 - m^2c^2}$; the momentum in Region 2 can be positive or negative, $p_2 = \pm\sqrt{(E - V_0)^2/c^2 - m^2c^2}$.

The constants B, C, D, and E can be obtained in terms of A from the four boundary conditions of the wave function and its derivative at $z = 0$ and $z = a$:

$$\psi_1(0) = \psi_2(0) \quad \Longrightarrow \quad A + B = C + D, \tag{12.882}$$

$$\frac{d\psi_1(0)}{dz} = \frac{d\psi_2(0)}{dz} \quad \Longrightarrow \quad A - B = \frac{p_2}{p_1}(C - D). \tag{12.883}$$

$$\psi_2(a) = \psi_3(a) \quad \Longrightarrow \quad Ce^{ip_2a/\hbar} + De^{-ip_2a/\hbar} = Ee^{ip_1a/\hbar}, \tag{12.884}$$

$$\frac{d\psi_2(a)}{dz} = \frac{d\psi_3(a)}{dz} \quad \Longrightarrow \quad Ce^{ip_2a/\hbar} - De^{-ip_2a/\hbar} = \frac{p_1}{p_2}Ee^{ip_1a/\hbar}. \tag{12.885}$$

Now, since the momenta in the $z \le 0$ and $z \ge a$ regions are equal, $p_1 = p_3$, the transmission coefficient is given by

$$T = \frac{p_3|E|^2}{p_1|A|^2} = \frac{|E^2|}{|A|^2}. \tag{12.886}$$

So, to find T, we would need to calculate the constant E first; this can be obtained by combining equations (12.882)–(12.885). First, an addition of (12.882) and (12.883) leads to

$$2A = C\left(1 + \frac{p_2}{p_1}\right) + D\left(1 - \frac{p_2}{p_1}\right), \tag{12.887}$$

the addition of (12.884) and (12.885) to

$$2Ce^{ip_2a/\hbar} = E\left(1 + \frac{p_1}{p_2}\right)e^{ip_1a/\hbar} \quad \Longrightarrow \quad C = \frac{E}{2}\left(1 + \frac{p_1}{p_2}\right)e^{i(p_1 - p_2)a/\hbar} \tag{12.888}$$

and the subtraction of (12.885) from (12.884) to

$$2De^{-ip_2a/\hbar} = E\left(1-\frac{p_1}{p_2}\right)e^{ip_1a/\hbar} \qquad \Longrightarrow \qquad D = \frac{E}{2}\left(1-\frac{p_1}{p_2}\right)e^{i(p_1+p_2)a/\hbar}. \tag{12.889}$$

Second, an insertion of (12.888)) and (12.889) into (12.887) yields

$$\begin{aligned} 2A &= \frac{E}{2}\left[\left(1+\frac{p_2}{p_1}\right)\left(1+\frac{p_1}{p_2}\right)e^{-ip_2a/\hbar} + \left(1-\frac{p_2}{p_1}\right)\left(1-\frac{p_1}{p_2}\right)e^{ip_2a/\hbar}\right]e^{ip_1a/\hbar} \\ &= \frac{E}{2p_1p_2}\left[(p_1+p_2)^2\,e^{-ip_2a/\hbar} - (p_1-p_2)^2\,e^{ip_2a/\hbar}\right]e^{ip_1a/\hbar}, \end{aligned} \tag{12.890}$$

which leads to

$$\begin{aligned} 2A &= \frac{E}{2p_1p_2}\left\{\left[(p_1+p_2)^2-(p_1-p_2)^2\right]\cos\left(\frac{p_2a}{\hbar}\right) - i\left[(p_1+p_2)^2+(p_1-p_2)^2\right]\sin\left(\frac{p_2a}{\hbar}\right)\right\}e^{ip_1a/\hbar} \\ &= \frac{E}{2p_1p_2}\left[4p_1p_2\cos\left(\frac{p_2a}{\hbar}\right) - 2i\left(p_1^2+p_2^2\right)\sin\left(\frac{p_2a}{\hbar}\right)\right]e^{ip_1a/\hbar}. \end{aligned} \tag{12.891}$$

Thus, the coefficient E is given by

$$E = 4p_1p_2Ae^{-ip_1a/\hbar}\left[4p_1p_2\cos\left(\frac{p_2a}{\hbar}\right) - 2i\left(p_1^2+p_2^2\right)\sin\left(\frac{p_2a}{\hbar}\right)\right]^{-1}. \tag{12.892}$$

Finally, an insertion of (12.892) into (12.886) leads at once to the transmission coefficient:

$$\begin{aligned} T &= \frac{|E|^2}{|A|^2} = \left|\frac{4p_1p_2e^{-ip_1a/\hbar}}{4p_1p_2\cos(p_2a/\hbar) - 2i\left(p_1^2+p_2^2\right)\sin(p_2a/\hbar)}\right|^2 \\ &= \frac{(4p_1p_2)^2}{(4p_1p_2)^2\cos^2(p_2a/\hbar) + 4\left(p_1^2+p_2^2\right)^2\sin^2(p_2a/\hbar)} \\ &= \frac{(4p_1p_2)^2}{(4p_1p_2)^2 + \left[-(4p_1p_2)^2 + 4\left(p_1^2-p_2^2\right)^2\right]\sin^2(p_2a/\hbar)}; \end{aligned} \tag{12.893}$$

we can reduce this expression to

$$T = \left[1+\frac{\left(p_1^2-p_2^2\right)^2}{4p_1^2p_2^2}\sin^2\left(p_2a/\hbar\right)\right]^{-1}. \tag{12.894}$$

Using the fact that $c^2p^2 = E^2 - m^2c^4$ and $c^2p_2^2 = (E-V_0)^2 - m^2c^4$, we have

$$\frac{\left(p_1^2-p_2^2\right)^2}{p_1^2p_2^2} = \frac{V_0^2(2E-V_0)^2}{4(E^2-m^2c^4)\left[(E-V_0)^2-m^2c^4\right]}, \tag{12.895}$$

and, hence, we can rewrite (12.894) as

$$T = \left[1+\frac{V_0^2(2E-V_0)^2}{4(E^2-m^2c^4)\left[(E-V_0)^2-m^2c^4\right]}\sin^2\left(\frac{a}{\hbar}\sqrt{(E-V_0)^2-m^2c^4}\right)\right]^{-1}. \tag{12.896}$$

This result is similar to the one we obtained in Section 4.5 using the Schrödinger equation.

Following the same method above that led to the transmission coefficient, we can show that the boundary conditions (12.882)–(12.885) yield the following expression for the reflection coefficient

$$R = \left[1 + \frac{4p_1^2p_2^2}{\left(p_1^2 - p_2^2\right)^2}\frac{1}{\sin^2(p_2a/\hbar)}\right]^{-1} = \frac{\sin^2(p_2a/\hbar)}{\frac{4p_1^2p_2^2}{(p_1^2-p_2^2)^2} + \sin^2(p_2a/\hbar)}. \tag{12.897}$$

From (12.894) and (12.897), we can easily see that both of R and T are smaller than 1: $R < 1$ amd $T < 1$. This was expected for the energy range $E > V_0 + mc^2$ where we have both reflected and transmitted waves. Moreover, by adding (12.894) and (12.897), we can verify that the reflected and transmission coefficients add up to unity:

$$R + T = \left[1 + \frac{4p_1^2p_2^2}{\left(p_1^2 - p_2^2\right)^2}\frac{1}{\sin^2(p_2a/\hbar)}\right]^{-1} + \left[1 + \frac{(p_1^2 - p_2^2)^2}{4p_1^2p_2^2}\sin^2(p_2a/\hbar)\right]^{-1} = 1. \tag{12.898}$$

An interesting case happens when $p_2a = n\pi$. As can be seen from (12.894) and (12.897), when $p_2a = n\pi$, we have total transmission: $R = 0$ and $T = 1$. This can be interpreted as follows: while the antiparticles created at $z = a$ travel left and annihilate the incident particles at $x = 0$, the particles created at $z = a$ will travel to the right into the $z > a$ region. Hence, nothing is reflected in the $z < 0$ region irrespective of the thickness of the barrier.

(b) When $V_0 > E + mc^2$ (i.e., $E < V_0 - mc^2$), the momentum in the $0 < z < a$ region is real and can be either positive or negative: $p_2 = \pm\sqrt{(E - V_0)^2 - m^2c^4}/c^2$. Since $E < V_0$ as displayed in right-side of Figure 12.8, the wave function (12.881) remains the same except that we need to replace ψ_2 by $\psi_2(z) = Ce^{p_2z/\hbar} + De^{-p_2z/\hbar}$:

$$\psi(z) = \begin{cases} \psi_1(z) = Ae^{ip_1z/\hbar} + Be^{-ip_1z/\hbar}, & z \le 0, \\ \psi_2(z) = Ce^{p_2z/\hbar} + De^{-p_2z/\hbar}, & 0 < z < a, \\ \psi_3(z) = Ee^{ip_1z/\hbar}, & a \ge a. \end{cases} \tag{12.899}$$

The four boundary conditions can be written as

$$\psi_1(0) = \psi_2(0) \quad \Longrightarrow \quad A + B = C + D, \tag{12.900}$$

$$\frac{d\psi_1(0)}{dz} = \frac{d\psi_2(0)}{dz} \quad \Longrightarrow \quad A - B = \frac{p_2}{ip_1}(C - D). \tag{12.901}$$

$$\psi_2(a) = \psi_3(a) \quad \Longrightarrow \quad Ce^{p_2a/\hbar} + De^{-p_2a/\hbar} = Ee^{ip_1a/\hbar}, \tag{12.902}$$

$$\frac{d\psi_2(a)}{dz} = \frac{d\psi_3(a)}{dz} \quad \Longrightarrow \quad Ce^{p_2a/\hbar} - De^{-p_2a/\hbar} = \frac{ip_1}{p_2}Ee^{ip_1a/\hbar}. \tag{12.903}$$

By adding (12.902) to (12.903), then subtracting (12.903) from (12.902) and then substituting the two resulting equations into (12.900) and (12.901), we can show that

$$\frac{B}{A} = -i\frac{p_1^2 + p_2^2}{p_1p_2}\sinh(p_2a/\hbar)\left[2\cosh(p_2a/\hbar) + i\frac{p_2^2 - p_1^2}{p_1p_2}\sinh(p_2a/\hbar)\right]^{-1}, \tag{12.904}$$

$$\frac{E}{A} = 2e^{-ip_1a}\left[2\cosh(p_2a/\hbar) + i\frac{p_2^2 - p_1^2}{p_1p_2}\sinh(p_2a/\hbar)\right]^{-1}. \tag{12.905}$$

Hence, the reflection and transmission coefficients are given by

$$R = \frac{|B|^2}{|A|^2} = \left(\frac{p_1^2 + p_2^2}{p_1 p_2}\right)^2 \sinh^2(p_2 a/\hbar) \left[4\cosh^2(p_2 a/\hbar) + \left(\frac{p_2^2 - p_1^2}{p_1 p_2}\right)^2 \sinh^2(p_2 a/\hbar)\right]^{-1}, \quad (12.906)$$

$$T = \frac{|E|^2}{|A|^2} = 4\left[4\cosh^2(p_2 a/\hbar) + \left(\frac{p_2^2 - p_1^2}{p_1 p_2}\right)^2 \sinh^2(p_2 a/\hbar)\right]^{-1}. \quad (12.907)$$

Since $\cosh^2(p_2 a/\hbar) = 1 + \sinh^2(p_2 a/\hbar)$, we can rewrite R in terms of T as

$$R = \left[1 + 4\left(\frac{p_1 p_2}{p_1^2 + p_2^2}\right)^2 \frac{1}{\sinh^2(p_2 a/\hbar)}\right]^{-1}, \quad (12.908)$$

$$T = \left[1 + \frac{1}{4}\left(\frac{p_1^2 + p_2^2}{p_1 p_2}\right)^2 \sinh^2(p_2 a/\hbar)\right]^{-1}. \quad (12.909)$$

The transmission into the $z > a$ region occurs by tunneling through the $0 < z < a$ barrier. Additionally, using (12.908) and (12.909), we see that both of R and T are less than 1 and that they add up to 1:

$$R < 1, \qquad T < 1 \qquad \text{and} \qquad R + T = 1. \quad (12.910)$$

Remarks

- If $E \ll V_0$, we can use the following approximations: $p_2 \sim V_0/c$, $(p_1^2+p_2^2)/(p_1 p_2) \sim V_0/(cp_1)$ and $\sinh(p_2 a/\hbar) \sim \sinh(V_0 a/\hbar c) \gg$, since $\sinh(V_0 a/\hbar c)$ is a very large quantity. Hence, the transmission coefficient (12.909) becomes asymptotically equal to

$$T \simeq \left[1 + \frac{cp_1}{4V_0}\sinh^2(V_0 a/\hbar c)\right]^{-1}. \quad (12.911)$$

  This shows that the transmission coefficient is small, but not zero. So, there is a finite tunneling through the classically forbidden region — the barrier — into the $z > a$ region. This can happen only due to the creation of particle-antiparticle pairs.

- When $E \simeq V_0$, the momentum becomes imaginary: $p_2 \sim imc$. We can then write: $\sinh(p_2 a/\hbar) \simeq \sin(mca/\hbar) \equiv \sin(p_2 a/\hbar)$. We can verify that the reflection and transmission coefficients (12.908) and (12.909) reduce to (12.897) and (12.894), respectively.

- If we take the classical limit $\hbar \to 0$, we have $\sinh(p_2 a/\hbar) \longrightarrow \infty$ and, hence the reflection and transmission coefficients (12.908) and (12.909) reduce to the classical result: $R \longrightarrow 1$ and $T \longrightarrow 0$, since when $E < V_0$, there can be no transmission; we have total reflection.

Problem 12.31
Using the matrices of β, α_x, α_y, α_z, and $\hat{\mathcal{P}}$, verify that the 4-current $j^\mu = \overline{\psi}\gamma^\mu\psi$ transforms as a vector under parity; i.e., show that $j^0 \longrightarrow j^0$ and $\vec{j} \longrightarrow -\vec{j}$.

Solution

The calculation of the transformation of j^0 under parity is straightforward; as shown in (12.553) and (12.547), since the parity operator is unitary,

$$\hat{\mathcal{P}}^\dagger\hat{\mathcal{P}} = \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix} = \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix} = I_4, \tag{12.912}$$

we have

$$j^0 \xrightarrow{\hat{\mathcal{P}}} \hat{\mathcal{P}}^\dagger j^0 \hat{\mathcal{P}} = c\,\psi^\dagger\left(\hat{\mathcal{P}}^\dagger\hat{\mathcal{P}}\right)\psi = c\psi^\dagger\psi = j^0. \tag{12.913}$$

As for the transformation of $\vec{j}$, we can easily infer it from Eq. (12.281); since $\vec{j} = c\psi^\dagger\vec{\alpha}\psi$, we can write

$$\vec{j} \xrightarrow{\hat{\mathcal{P}}} \hat{\mathcal{P}}^\dagger\vec{j}\hat{\mathcal{P}} = c\psi^\dagger\left(\hat{\mathcal{P}}^\dagger\,\vec{\alpha}\,\hat{\mathcal{P}}\right)\psi. \tag{12.914}$$

Let us now obtain the transformations of the three components j_x, j_y, j_z separately. Using the matrix of α_x listed in (12.163) and the matrix form of the parity operator as displayed in (12.547), we can obtain the transformation of the $x-$component of $\vec{j}$ as follows:

$$j_x \xrightarrow{\hat{\mathcal{P}}} \hat{\mathcal{P}}^\dagger j_x\hat{\mathcal{P}} = c\psi^\dagger\left(\hat{\mathcal{P}}^\dagger\,\alpha_x\,\hat{\mathcal{P}}\right)\psi, \tag{12.915}$$

where

$$\begin{aligned}\hat{\mathcal{P}}^\dagger\alpha_x\hat{\mathcal{P}} &= \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\\ &= \begin{pmatrix}0&0&0&-1\\0&0&-1&0\\0&-1&0&0\\-1&0&0&0\end{pmatrix}\\ &= -\alpha_x.\end{aligned} \tag{12.916}$$

Inserting (12.916) into (12.915), we see that the transformation of j_x is equal to $-jx$:

$$j_x \xrightarrow{\hat{\mathcal{P}}} \hat{\mathcal{P}}^\dagger j_x\hat{\mathcal{P}} = -c\psi^\dagger\left(\hat{\mathcal{P}}^\dagger\,\alpha_x\,\hat{\mathcal{P}}\right)\psi = -j_x. \tag{12.917}$$

Following the same method, we can easily obtain the transformations of α_y and α_z:

$$\begin{aligned}\hat{\mathcal{P}}^\dagger\alpha_y\hat{\mathcal{P}} &= \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0&0&0&-i\\0&0&i&0\\0&-i&0&0\\i&0&0&0\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\\ &= \begin{pmatrix}0&0&0&i\\0&0&-i&0\\0&i&0&0\\-i&0&0&0\end{pmatrix}\\ &= -\alpha_y,\end{aligned} \tag{12.918}$$

$$\begin{aligned}\hat{\mathcal{P}}^\dagger \alpha_z \hat{\mathcal{P}} &= \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0&0&1&0\\0&0&0&-1\\1&0&0&0\\0&-1&0&0\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\\ &= \begin{pmatrix}0&0&-1&0\\0&0&0&1\\-1&0&0&0\\0&1&0&0\end{pmatrix}\\ &= -\alpha_z, \end{aligned} \tag{12.919}$$

which lead to the transformation of j_y and j_z:

$$j_y \xrightarrow{\hat{\mathcal{P}}} \hat{\mathcal{P}}^\dagger j_y \hat{\mathcal{P}} = -c\psi^\dagger \left(\hat{\mathcal{P}}^\dagger \alpha_y \hat{\mathcal{P}}\right) \psi = -j_y, \tag{12.920}$$

$$j_z \xrightarrow{\hat{\mathcal{P}}} \hat{\mathcal{P}}^\dagger j_z \hat{\mathcal{P}} = -c\psi^\dagger \left(\hat{\mathcal{P}}^\dagger \alpha_z \hat{\mathcal{P}}\right) \psi = -j_z. \tag{12.921}$$

Finally, inserting (12.917), (12.920), and (12.921) into (12.914), we see that $\vec{j}$ transforms like a vector:

$$\vec{j} \xrightarrow{\hat{\mathcal{P}}} \hat{\mathcal{P}}^\dagger \vec{j} \hat{\mathcal{P}} = -c\psi^\dagger \left(\hat{\mathcal{P}}^\dagger \vec{\alpha} \hat{\mathcal{P}}\right) \psi = -\vec{j}. \tag{12.922}$$

Problem 12.32
Find the transformation of the right-chiral and left-chiral states ψ_R and ψ_L under parity by using the Weyl representation of the γ matrices (12.226).

Solution

Using the Weyl basis or the Weyl representation of the γ matrices (12.226), the matrix of the parity operator (12.547) becomes

$$\hat{\mathcal{P}} = \pm\gamma^0 = \pm\begin{pmatrix}0 & I_2\\ I_2 & 0\end{pmatrix} = \pm\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix}. \tag{12.923}$$

As shown in Eq. (12.425), the left-chiral states are stacked in the upper part of the Dirac spinor, while the right-chiral states in the lower part:

$$\psi_L = \begin{pmatrix}\phi_1\\ \phi_2\\ 0\\ 0\end{pmatrix}, \qquad \psi_R = \begin{pmatrix}0\\ 0\\ \phi_3\\ \phi_4\end{pmatrix}. \tag{12.924}$$

Using (12.923), we can show that the parity operator changes the helicity by transforming left-chiral states into right-chiral states and vice-versa:

$$\hat{\mathcal{P}}\psi_L = \pm\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix}\begin{pmatrix}\phi_1\\ \phi_2\\ 0\\ 0\end{pmatrix} = \pm\begin{pmatrix}0\\ 0\\ \phi_2\\ \phi_1\end{pmatrix} \equiv \pm\psi_R, \tag{12.925}$$

$$\hat{\mathcal{P}}\psi_R = \pm\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix}\begin{pmatrix}0\\0\\\phi_3\\\phi_4\end{pmatrix} = \pm\begin{pmatrix}\phi_4\\\phi_3\\0\\0\end{pmatrix} \equiv \pm\psi_L. \tag{12.926}$$

Hence, chirality flips under spatial inversion; that is, parity flips a left-chiral state into a right-chiral state and vice versa.

Problem 12.33
By taking $\vec{A} = 0$ in (12.462) and expanding the Dirac spinor ϕ_B up to order $1/c^2$, show that the nonrelativistic limit of the Dirac Hamiltonian contains the various correction terms: the relativistic correction to the kinetic energy as well as the spin-orbit coupling and the Darwin terms.

Solution
If we expand (12.462) up to order $1/c^2$ and, for simplicity, if we take $\vec{A} = 0$, we can write the Dirac spinor ϕ_B as

$$\phi_B = \frac{1}{2mc^2}\left(1 + \frac{E_{NR} - \hat{V}}{2mc^2}\right)^{-1} c\vec{\sigma}\cdot\hat{\vec{p}}\,\phi_A \simeq \frac{1}{2mc}\left(1 - \frac{E_{NR} - \hat{V}}{2mc^2}\right)\vec{\sigma}\cdot\hat{\vec{p}}\,\phi_A. \tag{12.927}$$

An insertion of this expression into the right-hand side of (12.460) leads to

$$\left(E_{NR} - \hat{V}\right)\phi_A \simeq \frac{1}{2m}\vec{\sigma}\cdot\hat{\vec{p}}\left(1 - \frac{E_{NR} - \hat{V}}{2mc^2}\right)\vec{\sigma}\cdot\hat{\vec{p}}\,\phi_A \tag{12.928}$$

or to

$$\left[\frac{1}{2m}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)^2 + \hat{V} - \frac{1}{4m^2c^2}\vec{\sigma}\cdot\hat{\vec{p}}\left(E_{NR} - \hat{V}\right)\vec{\sigma}\cdot\hat{\vec{p}}\right]\phi_A = E_{NR}\phi_A, \tag{12.929}$$

which, using the relation $\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)^2 = \hat{\vec{p}}^{\,2}$, reduces to

$$\left[\frac{\hat{\vec{p}}^{\,2}}{2m} + \hat{V} - \frac{E_{NR}\,\hat{\vec{p}}^{\,2}}{4m^2c^2} + \frac{\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)}{4m^2c^2}\right]\phi_A = E_{NR}\phi_A. \tag{12.930}$$

It is important to note that, taken separately, the 2–component spinors ϕ_A and ϕ_B are not normalized; however, the 4–component spinor $\psi = \begin{pmatrix}\phi_A\\\phi_B\end{pmatrix}$ can be normalized:

$$1 = \langle\psi \mid \psi\rangle = \begin{pmatrix}\phi_A^* & \phi_B^*\end{pmatrix}\begin{pmatrix}\phi_A\\\phi_B\end{pmatrix} = \langle\phi_A \mid \phi_A\rangle + \langle\phi_B \mid \phi_B\rangle, \tag{12.931}$$

where $\langle\phi_B \mid \phi_B\rangle$ can be obtained from (12.927):

$$\begin{aligned}\langle\phi_B \mid \phi_B\rangle &= \frac{1}{(2mc)^2}\langle\phi_A \mid \vec{\sigma}\cdot\hat{\vec{p}}\left(1 - \frac{E_{NR} - \hat{V}}{2mc^2}\right)\left(1 - \frac{E_{NR} - \hat{V}}{2mc^2}\right)\vec{\sigma}\cdot\hat{\vec{p}} \mid \phi_A\rangle \\ &\simeq \frac{1}{(2mc)^2}\langle\phi_A \mid \left(\vec{\sigma}\cdot\hat{\vec{p}}\right)^2 - 2(\vec{\sigma}\cdot\hat{\vec{p}})\,\frac{(E_{NR} - \hat{V})}{2mc^2}(\vec{\sigma}\cdot\hat{\vec{p}}) \mid \phi_A\rangle \\ &\simeq \langle\phi_A \mid \frac{\hat{\vec{p}}^{\,2}}{4m^2c^2} \mid \phi_A\rangle,\end{aligned} \tag{12.932}$$

where, in the last line, we have used the fact that $\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)^2 = \hat{\vec{p}}^{\,2}$; additionally, we have retained only terms up to $1/c^2$ is all expressions above. In the first line of (12.932), we have also used the fact that both of $\vec{\sigma}$ and $\hat{\vec{p}}$ are Hermitian and hence: $\left[\left(1-\frac{E_{NR}-\hat{V}}{2mc^2}\right)\vec{\sigma}\cdot\hat{\vec{p}}\mid\phi_A\rangle\right]^\dagger = \langle\phi_A\mid\vec{\sigma}\cdot\hat{\vec{p}}\left(1-\frac{E_{NR}-\hat{V}}{2mc^2}\right)$. Next, inserting (12.932) into (12.931), we obtain

$$\langle\psi\mid\psi\rangle = \langle\phi_A\mid 1+\frac{\hat{\vec{p}}^{\,2}}{4m^2c^2}\mid\phi_A\rangle = 1. \tag{12.933}$$

We may now introduce a 2–component spinor ϕ_N which is normalized and defined in terms of ϕ_A as follows

$$\mid\phi_N\rangle = \left(1+\frac{\hat{\vec{p}}}{8m^2c^2}\right)\mid\phi_A\rangle \Longrightarrow \mid\phi_A\rangle = \left(1+\frac{\hat{\vec{p}}}{8m^2c^2}\right)^{-1}\mid\phi_N\rangle \simeq \left(1-\frac{\hat{\vec{p}}^{\,2}}{8m^2c^2}\right)\mid\phi_N\rangle. \tag{12.934}$$

Note that the spinor $\mid\phi_N\rangle$ in equation (12.934) does satisfy (12.933); namely

$$\begin{aligned}\langle\phi_N\mid\phi_N\rangle &= \langle\phi_A\mid\left(1+\frac{\hat{\vec{p}}}{8m^2c^2}\right)\left(1+\frac{\hat{\vec{p}}}{8m^2c^2}\right)\mid\phi_A\rangle\\ &\simeq \langle\phi_A\mid 1+\frac{\hat{\vec{p}}}{4m^2c^2}\mid\phi_A\rangle = 1,\end{aligned} \tag{12.935}$$

where, in the second line, we have kept only terms up to $1/c^2$.

On the one hand, substituting (12.934) into (12.930) and keeping only terms up to $1/c^2$, we obtain

$$\begin{aligned}&\langle\phi_A\mid\hat{H}_{NR}\mid\phi_A\rangle\\ &= \langle\phi_N\mid\vec{\sigma}\cdot\hat{\vec{p}}\left(1-\frac{\hat{\vec{p}}^{\,2}}{8m^2c^2}\right)\left[\frac{\hat{\vec{p}}^{\,2}}{2m}+\hat{V}-\frac{E_{NR}\,\hat{\vec{p}}^{\,2}}{4m^2c^2}+\frac{\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)}{4m^2c^2}\right]\vec{\sigma}\cdot\hat{\vec{p}}\left(1-\frac{\hat{\vec{p}}^{\,2}}{8m^2c^2}\right)\mid\phi_N\rangle\\ &\simeq \langle\phi_N\mid\vec{\sigma}\cdot\hat{\vec{p}}\left(1-\frac{\hat{\vec{p}}^{\,2}}{8m^2c^2}\right)\left[\frac{\hat{\vec{p}}^{\,2}}{2m}+\hat{V}-\frac{\hat{\vec{p}}^{\,4}}{16m^3c^2}-\frac{E_{NR}\,\hat{\vec{p}}^{\,2}}{4m^2c^2}+\frac{\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)}{4m^2c^2}-\frac{\hat{V}\hat{\vec{p}}^{\,2}}{8m^2c^2}\right]\mid\phi_N\rangle\\ &\simeq \langle\phi_N\mid\left[\frac{\hat{\vec{p}}^{\,2}}{2m}+\hat{V}-\frac{\hat{\vec{p}}^{\,4}}{8m^3c^2}-\frac{E_{NR}\,\hat{\vec{p}}^{\,2}}{4m^2c^2}+\frac{\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)}{4m^2c^2}-\frac{\left(\hat{V}\hat{\vec{p}}^{\,2}+\hat{\vec{p}}^{\,2}\hat{V}\right)}{8m^2c^2}\right]\mid\phi_N\rangle.\end{aligned} \tag{12.936}$$

On the other hand, using (12.934) and keeping only terms up to $1/c^2$, we can write

$$\begin{aligned}E_{NR}\langle\phi_A\mid\phi_A\rangle &= \langle\phi_N\mid\left(1-\frac{\hat{\vec{p}}}{8m^2c^2}\right)\left(1-\frac{\hat{\vec{p}}}{8m^2c^2}\right)\mid\phi_N\rangle\\ &\simeq \langle\phi_N\mid E_{NR}-\frac{E_{NR}\,\hat{\vec{p}}^{\,2}}{4m^2c^2}\mid\phi_N\rangle,\end{aligned} \tag{12.937}$$

Now, equating (12.936) and (12.937) and canceling the term $(E_{NR}\,\hat{\vec{p}}^{\,2})/(4m^2c^2)$ on both sides of the equation, we obtain

$$\left[\frac{\hat{\vec{p}}^{\,2}}{2m}+\hat{V}-\frac{\hat{\vec{p}}^{\,4}}{8m^3c^2}+\frac{\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)}{4m^2c^2}-\frac{\left(\hat{V}\hat{\vec{p}}^{\,2}+\hat{\vec{p}}^{\,2}\hat{V}\right)}{8m^2c^2}\right]\mid\phi_N\rangle = E_{NR}\mid\phi_N\rangle. \tag{12.938}$$

Hence, we can write the non-relativistic limit of the Dirac Hamiltonian as follows

$$\hat{H}_{NR} = \frac{\hat{\vec{p}}^{\,2}}{2m} + \hat{V} - \frac{\hat{\vec{p}}^{\,4}}{8m^3c^2} + \frac{\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)}{4m^2c^2} - \frac{(\hat{V}\hat{\vec{p}}^{\,2} + \hat{\vec{p}}^{\,2}\hat{V})}{8m^2c^2}\,. \tag{12.939}$$

Let us now work on the simplification the term $\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)$. Since $\hat{\vec{p}} = -i\hbar\vec{\nabla}$, we have

$$\begin{aligned}
\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right) &= -i\hbar\left(\vec{\sigma}\cdot\vec{\nabla}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right) \\
&= -i\hbar\left(\vec{\sigma}\cdot(\vec{\nabla}\hat{V})\right)\left(\vec{\sigma}\cdot\hat{\vec{p}}\right) - i\hbar\hat{V}\left(\vec{\sigma}\cdot\vec{\nabla}\right)\left(\vec{\sigma}\cdot\hat{\vec{p}}\right) \\
&\equiv -i\hbar\left(\vec{\sigma}\cdot(\vec{\nabla}\hat{V})\right)\left(\vec{\sigma}\cdot\hat{\vec{p}}\right) + \hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\left(\vec{\sigma}\cdot\hat{\vec{p}}\right) \\
&= -i\hbar\left(\vec{\sigma}\cdot(\vec{\nabla}\hat{V})\right)\left(\vec{\sigma}\cdot\hat{\vec{p}}\right) + \hat{V}\hat{\vec{p}}^{\,2},
\end{aligned} \tag{12.940}$$

where, in the last line, we have used the fact that $\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\left(\vec{\sigma}\cdot\hat{\vec{p}}\right) \equiv \left(\vec{\sigma}\cdot\hat{\vec{p}}\right)^2 = \hat{\vec{p}}^{\,2}$. Now, using the relation (12.466), we can write the first term in (12.940) as follows

$$\left(\vec{\sigma}\cdot(\vec{\nabla}\hat{V})\right)\left(\vec{\sigma}\cdot\hat{\vec{p}}\right) = (\vec{\nabla}\hat{V})\cdot\hat{\vec{p}} + i\vec{\sigma}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right), \tag{12.941}$$

which, when inserted back into (12.940, leads to

$$\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right) = -i\hbar(\vec{\nabla}\hat{V})\cdot\hat{\vec{p}} + \hbar\vec{\sigma}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right) + \hat{V}\hat{\vec{p}}^{\,2}. \tag{12.942}$$

Next, using (12.942), we can reduce the last two terms in (12.939) to

$$\begin{aligned}
\frac{\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)}{4m^2c^2} - \frac{(\hat{V}\hat{\vec{p}}^{\,2} + \hat{\vec{p}}^{\,2}\hat{V})}{8m^2c^2} &= -iq\hbar\frac{(\vec{\nabla}\hat{V})\cdot\hat{\vec{p}}}{4m^2c^2} + q\hbar\frac{\vec{\sigma}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right)}{4m^2c^2} + \frac{\hat{V}\hat{\vec{p}}^{\,2}}{4m^2c^2} - \frac{(\hat{V}\hat{\vec{p}}^{\,2} + \hat{\vec{p}}^{\,2}\hat{V})}{8m^2c^2} \\
&= -i\hbar\frac{(\vec{\nabla}\hat{V})\cdot\hat{\vec{p}}}{4m^2c^2} + q\hbar\frac{\vec{\sigma}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right)}{4m^2c^2} + \frac{(\hat{V}\hat{\vec{p}}^{\,2} - \hat{\vec{p}}^{\,2}\hat{V})}{8m^2c^2}.
\end{aligned} \tag{12.943}$$

To simplify $(\hat{V}\hat{\vec{p}}^{\,2} - \hat{\vec{p}}^{\,2}\hat{V})$ and invoking $\nabla^2(fg) = (\nabla^2 f)g + 2(\nabla f)\cdot(\nabla g) + f(\nabla^2 g)$, we can write

$$\nabla^2(\hat{V}\phi) = \vec{\nabla}\cdot\left((\vec{\nabla}\,\hat{V})\phi + \hat{V}(\vec{\nabla}\phi)\right) = (\nabla^2\hat{V})\phi + 2(\vec{\nabla}\hat{V})\cdot(\vec{\nabla}\phi) + \hat{V}\nabla^2\phi, \tag{12.944}$$

which, after multiplying it by $-\hbar^2$ and carrying out some rearrangements, leads to

$$-\hbar^2(\nabla^2\hat{V})\phi + \hbar^2\hat{V}\nabla^2\phi = -\hbar^2(\nabla^2\hat{V}) - 2i\hbar(\vec{\nabla}\hat{V})\cdot(-i\hbar\vec{\nabla}\phi)\,, \tag{12.945}$$

or to

$$\hat{V}\hat{\vec{p}}^{\,2} - \hat{\vec{p}}^{\,2}\hat{V} = \hbar^2(\nabla^2\hat{V}) + 2i\hbar(\vec{\nabla}\hat{V})\cdot\hat{\vec{p}}, \tag{12.946}$$

where we have used the fact that $-i\hbar\vec{\nabla} = \hat{\vec{p}}$ and $-\hbar^2\nabla^2 = \hat{\vec{p}}^{\,2}$. Inserting (12.946) into (12.943)

and performing some simplifications, we obtain

$$\begin{aligned}
\frac{\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)}{4m^2c^2} - \frac{(\hat{V}\hat{\vec{p}}^{\,2}+\hat{\vec{p}}^{\,2}\hat{V})}{8m^2c^2} &= -i\hbar\frac{(\vec{\nabla}\hat{V})\cdot\hat{\vec{p}}}{4m^2c^2} + \hbar\frac{\vec{\sigma}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right)}{4m^2c^2} + \frac{(\hat{V}\hat{\vec{p}}^{\,2}-\hat{\vec{p}}^{\,2}\hat{V})}{8m^2c^2} \\
&= -i\hbar\frac{(\vec{\nabla}\hat{V})\cdot\hat{\vec{p}}}{4m^2c^2} + \hbar\frac{\vec{\sigma}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right)}{4m^2c^2} + \frac{\hbar^2}{8m^2c^2}(\nabla^2\hat{V}) + i\hbar\frac{(\vec{\nabla}\hat{V})\cdot\hat{\vec{p}}}{4m^2c^2} \\
&= \frac{\hbar}{4m^2c^2}\vec{\sigma}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right) + \frac{\hbar^2}{8m^2c^2}(\nabla^2\hat{V}) \\
&= \frac{1}{2m^2c^2}\hat{\vec{S}}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right) + \frac{\hbar^2}{8m^2c^2}(\nabla^2\hat{V}), \qquad (12.947)
\end{aligned}$$

where, in the last line, we have used the spin expression $\hat{\vec{S}} = (\hbar/2)\vec{\sigma}$.

Finally, after substituting (12.947) into the non-relativistic Dirac Hamiltonian (12.939), we obtain

$$\begin{aligned}
\hat{H}_{NR} &= \frac{\hat{\vec{p}}^{\,2}}{2m} + \hat{V} - \frac{\hat{\vec{p}}^{\,4}}{8m^3c^2} + \frac{\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)\hat{V}\left(\vec{\sigma}\cdot\hat{\vec{p}}\right)}{4m^2c^2} - \frac{(\hat{V}\hat{\vec{p}}^{\,2}+\hat{\vec{p}}^{\,2}\hat{V})}{8m^2c^2} \\
&= \frac{\hat{\vec{p}}^{\,2}}{2m} + \hat{V} - \frac{\hat{\vec{p}}^{\,4}}{8m^3c^2} + \frac{1}{2m^2c^2}\hat{\vec{S}}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right) + \frac{\hbar^2}{8m^2c^2}(\nabla^2\hat{V}) \qquad (12.948)
\end{aligned}$$

which can be rewritten as

$$\boxed{\hat{H}_{NR} = \frac{\hat{\vec{p}}^{\,2}}{2m} + \hat{V} - \frac{\hat{\vec{p}}^{\,4}}{8m^3c^2} + \hat{H}_{SO} + \hat{H}_{Darwin}, \quad \hat{H}_{SO} = \frac{1}{2m^2c^2}\hat{\vec{S}}\cdot\left((\vec{\nabla}\hat{V})\times\hat{\vec{p}}\right), \quad \hat{H}_{Darwin} = \frac{\hbar^2}{8m^2c^2}(\nabla^2\hat{V}),} \qquad (12.949)$$

where $-\hat{\vec{p}}^{\,4}/(8m^3c^2)$ represents the relativistic correction to the kinetic energy, $\hat{H}_{SO}$ the spin-orbit coupling term, and $\hat{H}_{Darwin}$ the Darwin term.

Problem 12.34

Consider the particle and anti-particle helicity eigenstates (12.408) and (12.409).

(a) Find the normalization constant N of these eigenstates.

(b) Show that these helicity eigenstates are orthogonal.

Solution

(a) To normalize the eigenstates (12.408) and (12.409), we simply need to normalize any one of $|u\rangle_+$, $|u\rangle_-$, $|v\rangle_+$, $|v\rangle_-$; they all lead to the same value of N. We need to keep in mind that $\langle\,\psi_\pm\,|\,\psi_\pm\,\rangle = {}_\pm\langle u\,|\,u\rangle_\pm$ and $\langle\,\Phi_\pm\,|\,\Phi_\pm\,\rangle = {}_\pm\langle v\,|\,v\rangle_\pm$. So, the normalization of $|\,\psi_\pm\rangle$ and $|\,\Phi_\pm\rangle$ reduce to normalizing $|\,u\rangle_\pm$ and $|\,v\rangle_\pm$, respectively.

To find N, and to ensure adherence with causality[84], we can select a normalization condition so that the particle and anti-particle helicity spinors $|u\rangle_+$, $|u\rangle_-$, $|v\rangle_+$, $|v\rangle_-$ transform like the *time-like* component of a Lorentz 4-vector with the requirement of having $2E$ particles per unit volume:

$$_+\langle u|u\rangle_+ = {}_-\langle u|u\rangle_- = {}_+\langle v|v\rangle_+ = {}_-\langle v|v\rangle_- = 2E. \qquad (12.950)$$

[84] As discussed in Appendix E, *time-like* events can be connected by causal, real physical signals.

Let us illustrate the calculations on $|u\rangle_+$. After taking the Hermitian adjoint of $|u\rangle_-$ from equation (12.408) and making use of the fact that p and E are Hermitian, we have

$$
\begin{aligned}
{}_+\langle u|u\rangle_+ &= |N|^2 \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & \sin\left(\frac{\theta}{2}\right) e^{-i\phi} & \frac{cp\cos\left(\frac{\theta}{2}\right)}{E+mc^2} & \frac{cp\sin\left(\frac{\theta}{2}\right)}{E+mc^2} e^{-i\phi} \end{pmatrix} \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) e^{i\phi} \\ \frac{cp\cos\left(\frac{\theta}{2}\right)}{E+mc^2} \\ \frac{cp\sin\left(\frac{\theta}{2}\right)}{E+mc^2} e^{i\phi} \end{pmatrix} \\
&= |N|^2 \left[\cos^2\left(\frac{\theta}{2}\right) + \sin^2\left(\frac{\theta}{2}\right) + \frac{c^2p^2}{(E+mc^2)^2}\cos^2\left(\frac{\theta}{2}\right) + \frac{c^2p^2}{(E+mc^2)^2}\sin^2\left(\frac{\theta}{2}\right)\right] \\
&= |N|^2 \left[1 + \frac{c^2p^2}{(E+mc^2)^2}\right] \equiv \frac{(E+mc^2)^2 + c^2p^2}{(E+mc^2)^2} \\
&= |N|^2 \frac{(E+mc^2)^2 + E^2 - m^2c^4}{(E+mc^2)^2} \\
&= |N|^2 \frac{2E}{E+mc^2},
\end{aligned}
\tag{12.951}
$$

where we have used the relation $c^2p^2 = E^2 - m^2c^4$ in the fourth line. Combining (12.951) with (12.950), we can obtain easily the value of the normalization constant

$$
\boxed{N = \sqrt{E+mc^2}.} \tag{12.952}
$$

(b) Using Eq. (12.408), we can easily show that the particle helicity states are orthogonal:

$$
\begin{aligned}
{}_+\langle u|u\rangle_- &= |N|^2 \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & \sin\left(\frac{\theta}{2}\right) e^{-i\phi} & \frac{cp\cos\left(\frac{\theta}{2}\right)}{E+mc^2} & \frac{cp\sin\left(\frac{\theta}{2}\right)}{E+mc^2} e^{-i\phi} \end{pmatrix} \begin{pmatrix} \sin\left(\frac{\theta}{2}\right) \\ -\cos\left(\frac{\theta}{2}\right) e^{i\phi} \\ -\frac{cp\sin\left(\frac{\theta}{2}\right)}{E+mc^2} \\ \frac{cp\cos\left(\frac{\theta}{2}\right)}{E+mc^2} \end{pmatrix} \\
&= |N|^2 \left[1 - 1 - \frac{c^2p^2}{(E+mc^2)^2} + \frac{c^2p^2}{(E+mc^2)^2}\right] \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) \\
&= 0,
\end{aligned}
\tag{12.953}
$$

Similarly, we can easily show that the anti-particle helicity states (12.409) are orthogonal:

$$
\begin{aligned}
{}_+\langle v|v\rangle_- &= |N|^2 \begin{pmatrix} -\dfrac{cp\sin\left(\frac{\theta}{2}\right)}{E+mc^2} & \dfrac{cp\cos\left(\frac{\theta}{2}\right)}{E+mc^2}\, e^{-i\phi} & \sin\left(\frac{\theta}{2}\right) & -\cos\left(\frac{\theta}{2}\right) e^{-i\phi} \end{pmatrix} \begin{pmatrix} \frac{cp\cos\left(\frac{\theta}{2}\right)}{E+mc^2} \\ \frac{cp\sin\left(\frac{\theta}{2}\right)}{E+mc^2}\, e^{i\phi} \\ \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) e^{i\phi} \end{pmatrix} \\
&= |N|^2 \left[-\frac{c^2p^2}{(E+mc^2)^2} + \frac{c^2p^2}{(E+mc^2)^2} + 1 - 1 \right] \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) \\
&= 0, \qquad (12.954)
\end{aligned}
$$

Combining Eqs. (12.950), (12.953) and (12.954), we see that the particle and anti-particle helicity eigenstates (12.408) and (12.409) are normalized to $2E$ and orthogonal:

$$\langle \psi_\pm \mid \psi_\pm \rangle \equiv {}_\pm\langle u \mid u\rangle_\pm = 2E, \qquad \langle \Phi_\pm \mid \Phi_\pm \rangle \equiv {}_\pm\langle v \mid v\rangle_\pm = 2E, \qquad (12.955)$$

$$\langle \psi_\pm \mid \psi_\mp \rangle \equiv {}_\pm\langle u \mid u\rangle_\mp = 0, \qquad \langle \Phi_\pm \mid \Phi_\mp \rangle \equiv {}_\pm\langle v \mid v\rangle_\mp = 0. \qquad (12.956)$$

Problem 12.35

Consider a spin-1/2 free particle. As we shown in the text, the Hamiltonian, momentum, and helicity operators mutually commute. Find the joint eigenstates to $\hat{H}$, $\hat{\vec{p}}$, and $\widehat{\mathcal{H}}$ for the following two separate cases where the particle moves

(a) along the $z-$axis, and

(b) along an arbitrary direction $(\theta,\ \phi)$.

Find these joint eigenstates for both particles and their antiparticles.

Solution

(a) As shown in the text, *in the particular case of a free particle moving along the z–direction* (i.e., $\hat{n} = (0,\ 0,\ \pm 1) \equiv \pm\hat{k}$), *the Dirac Hamiltonian commutes with* $\hat{S}_z \equiv \hat{\vec{S}} \cdot \hat{n}$; their joint eigenstates are given by the spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$ that were derived in Eq (12.336):

$$u_1 = N \begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \end{pmatrix}, \qquad u_2 = N \begin{pmatrix} 0 \\ 1 \\ \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix}, \qquad v_1 = N \begin{pmatrix} \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix}, \qquad v_2 = N \begin{pmatrix} \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \\ 1 \\ 0 \end{pmatrix}. \qquad (12.957)$$

In equations (12.392), (12.393), and (12.400), we have shown that these spinors are eigenstates of the Hamiltonian and the spin angular momentum.

Additionally, we can show that the Dirac spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$ are also eigenstates of the helicity operator. For this, invoking the following helicity matrix that was derived in

Chapter 12:

$$\widehat{\mathcal{H}} = \frac{\hbar}{2}\begin{pmatrix} \vec{\sigma}\cdot\hat{n} & 0 \\ 0 & \vec{\sigma}\cdot\hat{n} \end{pmatrix} = \frac{\hbar}{2|\vec{p}|}\begin{pmatrix} \hat{p}_z & \hat{p}_x - i\hat{p}_y & 0 & 0 \\ \hat{p}_x + i\hat{p}_y & -\hat{p}_z & 0 & 0 \\ 0 & 0 & \hat{p}_z & \hat{p}_x - i\hat{p}_z \\ 0 & 0 & \hat{p}_x + i\hat{p}_y & -\hat{p}_z \end{pmatrix}, \tag{12.958}$$

and inserting $p_x = 0$, $p_y = 0$, and $p_z = \pm p$ (with $p > 0$) into (12.957) and (12.958), we obtain

$$\boxed{|u_1\rangle = N\begin{pmatrix} 1 \\ 0 \\ \frac{\pm cp}{E^{(\cdot)}+mc^2} \\ 0 \end{pmatrix}, \quad |u_2\rangle = N\begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{\mp cp}{E^{(\cdot)}+mc^2} \end{pmatrix}, \quad |v_1\rangle = N\begin{pmatrix} 0 \\ \frac{\mp cp}{E^{(\cdot)}+mc^2} \\ 0 \\ 1 \end{pmatrix}, \quad |v_2\rangle = N\begin{pmatrix} \frac{\pm cp}{E^{(\cdot)}+mc^2} \\ 0 \\ 1 \\ 0 \end{pmatrix},} \tag{12.959}$$

and

$$\widehat{\mathcal{H}} = \frac{\hbar}{2p}\begin{pmatrix} p & 0 & 0 & 0 \\ 0 & -p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & -p \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{12.960}$$

Applying the helicity operator (12.960) on the spinors (12.959), we can see at once that the spinors $|u_1\rangle$, $|u_2\rangle$, $|v_1\rangle$, $|v_2\rangle$ are indeed eigenstates of the operator $\widehat{\mathcal{H}}$ with eigenvalues $\pm\hbar/2$:

$$\boxed{\widehat{\mathcal{H}}|u_1\rangle = \tfrac{\hbar}{2}|u_1\rangle, \quad \widehat{\mathcal{H}}|u_2\rangle = -\tfrac{\hbar}{2}|u_2\rangle, \quad \widehat{\mathcal{H}}|v_1\rangle = -\tfrac{\hbar}{2}|v_1\rangle, \quad \widehat{\mathcal{H}}|v_2\rangle = \tfrac{\hbar}{2}|v_2\rangle.} \tag{12.96}$$

As an illustration, the first relation was obtained as follow

$$\widehat{\mathcal{H}}|u_1\rangle = N\frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ \frac{\pm cp}{E^{(\cdot)}+mc^2} \\ 0 \end{pmatrix} = N\frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \\ \frac{\pm cp}{E^{(\cdot)}+mc^2} \\ 0 \end{pmatrix} \equiv \frac{\hbar}{2}|u_1\rangle, \tag{12.962}$$

(b) Now consider the case where the particle is moving in an arbitrary direction (θ, ϕ). Using the components of the momentum in spherical coordinates $p_x = p\sin\theta\cos\phi$, $p_y = p\sin\theta\sin\phi$, $p_z = p\cos\theta$, we can easily infer the unit vector along the momentum's direction:

$$\hat{n} = \frac{\vec{p}}{p} = \left(\frac{p_x}{p}, \frac{p_y}{p}, \frac{p_z}{p}\right) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \equiv (n_x, n_y, n_z), \tag{12.963}$$

which, when combined with the Pauli matrices, we obtain

$$\begin{aligned} \vec{\sigma}\cdot\hat{n} &= \sigma_x\frac{p_x}{p} + \sigma_y\frac{p_y}{p} + \sigma_z\frac{p_z}{p} = \frac{1}{p}\begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \sin\theta\cos\phi - i\sin\theta\sin\phi \\ \sin\theta\cos\phi + i\sin\theta\sin\phi & -\cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}. \end{aligned} \tag{12.964}$$

If we insert (12.964) into (12.958), we obtain the helicity operator

$$\widehat{\mathcal{H}} = \frac{\hbar}{2}\begin{pmatrix} \vec{\sigma}\cdot\hat{n} & 0 \\ 0 & \vec{\sigma}\cdot\hat{n} \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta & 0 & 0 \\ e^{i\phi}\sin\theta & -\cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & e^{-i\phi}\sin\theta \\ 0 & 0 & e^{i\phi}\sin\theta & -\cos\theta \end{pmatrix}. \tag{12.965}$$

Now, inserting $p_x = p\sin\theta\cos\phi$, $p_y = p\sin\theta\sin\phi$, $p_z = p\cos\theta$ into (12.168), we can write the Dirac Hamiltonian in spherical coordinates as follows

$$\hat{H} = cp\begin{pmatrix} mc/p & 0 & \cos\theta & e^{-i\phi}\sin\theta \\ 0 & mc/p & e^{i\phi}\sin\theta & -\cos\theta \\ \cos\theta & e^{-i\phi}\sin\theta & -mc/p & 0 \\ e^{i\phi}\sin\theta & -\cos\theta & 0 & -mc/p \end{pmatrix}. \tag{12.966}$$

We can easily show that the helicity and Hamiltonian operators (12.965) and (12.966) commute (see Problem 12.37 on Page 888 for a proof): $\left[\hat{H},\widehat{\mathcal{H}}\right] = 0$. Hence, they can have a set of joint eigenstates. In what follows, we are going to find these states that are solutions to the Dirac equation and are also eigenstates to the helicity operator. The eigenvalue problem for right-handed and left-handed helicity states can be inferred from (12.961):

$$\widehat{\mathcal{H}}|u\rangle_\pm = \pm\frac{\hbar}{2}|u\rangle_\pm, \qquad \widehat{\mathcal{H}}|v\rangle_\pm = \mp\frac{\hbar}{2}|v\rangle_\pm, \tag{12.967}$$

where $|u\rangle_\pm$ are right-handed and left-handed helicity states, respectively, for particles, while $|v\rangle_\pm$ are the respective right-handed and left-handed helicity states for antiparticles. Notice that $|u\rangle_\pm$ and $|v\rangle_\pm$ are four-component spinors. We can write $|u\rangle_\pm$ and $|v\rangle_\pm$ in terms of two-component spinors $|u_A\rangle_\pm$, $|u_B\rangle_\pm$, $|v_A\rangle_\pm$, $|v_B\rangle_\pm$ as follows

$$|u\rangle_\pm = \begin{pmatrix} |u_A\rangle_\pm \\ |u_B\rangle_\pm \end{pmatrix}, \qquad |v\rangle_\pm = \begin{pmatrix} |v_B\rangle_\pm \\ |v_A\rangle_\pm \end{pmatrix}. \tag{12.968}$$

In what follows, we are going to focus on determining the helicity eigenstates for particles and antiparticles.

Helicity eigenstates for particles

Let us now deal with the determination of the helicity eigenstates for particles; we will deal with the helicity eigenstates for antiparticles after that. So, combining (12.965), (12.967), and (12.968), we obtain

$$\widehat{\mathcal{H}}|u\rangle_\pm = \pm\frac{\hbar}{2}|u\rangle_\pm \quad\Longrightarrow\quad \begin{pmatrix} \vec{\sigma}\cdot\hat{n} & 0 \\ 0 & \vec{\sigma}\cdot\hat{n} \end{pmatrix}\begin{pmatrix} |u_A\rangle_\pm \\ |u_B\rangle_\pm \end{pmatrix} = \pm\begin{pmatrix} |u_A\rangle_\pm \\ |u_B\rangle_\pm \end{pmatrix}. \tag{12.969}$$

This equation leads to

$$\vec{\sigma}\cdot\hat{n}\ |u_A\rangle_\pm = \pm|u_A\rangle_\pm, \tag{12.970}$$

$$\vec{\sigma}\cdot\hat{n}\ |u_B\rangle_\pm = \pm|u_B\rangle_\pm. \tag{12.971}$$

Taking $|u_A\rangle_\pm = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ where a_1 and a_2 are complex numbers and using (12.964), we have

$$\vec{\sigma}\cdot\hat{n}\,|u_A\rangle_\pm = \pm|u_A\rangle_\pm \Longrightarrow \begin{pmatrix} \cos\theta & \sin\theta e^{-i\varphi} \\ \sin\theta e^{i\varphi} & -\cos\theta \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \pm\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \tag{12.972}$$

First, let us consider the positive helicity the positive helicity case: $\vec{\sigma}\cdot\hat{n}\,|u_A\rangle_+ = |u_A\rangle_+$; hence, (12.972) leads to

$$a_1\cos\theta + a_2 e^{-i\varphi}\sin\theta = a_1 \Longrightarrow a_1(1-\cos\theta) = a_2 e^{-i\varphi}\sin\theta \Longrightarrow a_2 = a_1\tan(\theta/2)\,e^{i\varphi}, \tag{12.973}$$

where we have used the facts that $(1-\cos\theta) = 2\sin^2(\theta/2)$ and $\sin\theta = 2\sin(\theta/2)\cos(\theta/2)$. Inserting these values of a_1 and a_2 into $|u_A\rangle_+ = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, we have

$$|u_A\rangle_+ = a_1\begin{pmatrix} 1 \\ \tan(\theta/2)\,e^{i\varphi} \end{pmatrix} \quad \text{or} \quad |u_A\rangle_+ = a_2\begin{pmatrix} \cot(\theta/2)\,e^{-i\varphi} \\ 1 \end{pmatrix}. \tag{12.974}$$

Now, we can obtain the numbers a_1 and a_2 by normalizing $|u_A\rangle_+$ to N^2:

$$N^2 = {}_+\langle u_A\,|\,u_A\rangle_+ = |a_1|^2\begin{pmatrix} 1 & \tan(\theta/2)\,e^{-i\varphi} \end{pmatrix}\begin{pmatrix} 1 \\ \tan(\theta/2)\,e^{i\varphi} \end{pmatrix} = |a_1|^2(1+\tan^2(\theta/2)) \equiv \frac{|a_1|^2}{\cos^2(\theta/2)}. \tag{12.975}$$

This yields $a_1 = N\cos(\theta/2)$ up to a phase factor which we are going to ignore as it will not affect the physics; this phase factor will cancel out when we deal with mean values. So, inserting a_1 into (12.974), we obtain

$$|u_A\rangle_+ = N\begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2)\,e^{i\varphi} \end{pmatrix} \quad \text{or} \quad |u_A\rangle_+ = N\begin{pmatrix} \cos(\theta/2)\,e^{-i\varphi} \\ \sin(\theta/2) \end{pmatrix}. \tag{12.976}$$

Throughout this subsection, we will be using the normalization constant $N = \sqrt{E^{(+)} + mc^2}$ that we was derived in Chapter 12 (see Eq. (12.722)).

Second, let us consider the negative helicity: $\vec{\sigma}\cdot\hat{n}\,|u_A\rangle_- = -|u_A\rangle_-$; hence, (12.972) leads to

$$a_1\cos\theta + a_2 e^{-i\varphi}\sin\theta = -a_1 \Longrightarrow a_1(1+\cos\theta) = -a_2 e^{-i\varphi}\sin\theta \Longrightarrow a_2 = -a_1\cot(\theta/2)\,e^{i\varphi}, \tag{12.977}$$

where we have used the fact that $(1+\cos\theta) = 2\cos^2(\theta/2)$ and $\sin\theta = 2\sin(\theta/2)\cos(\theta/2)$. Substituting the above values of a_1 and a_2 into $|u_A\rangle_- = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, we obtain

$$|u_A\rangle_- = a_1\begin{pmatrix} 1 \\ -\cot(\theta/2)\,e^{i\varphi} \end{pmatrix} \quad \text{or} \quad |u_A\rangle_- = a_2\begin{pmatrix} -\tan(\theta/2)\,e^{-i\varphi} \\ 1 \end{pmatrix}. \tag{12.978}$$

We now can obtain the numbers a_1 and a_2 by normalizing $|u_A\rangle_-$ to N^2 up to a phase factor to be ignored:

$$N^2 = {}_-\langle u_A\,|\,u_A\rangle_- = |a_1|^2\begin{pmatrix} 1 & -\cot(\theta/2)\,e^{-i\varphi} \end{pmatrix}\begin{pmatrix} 1 \\ -\cot(\theta/2)\,e^{i\varphi} \end{pmatrix} = |a_1|^2(1+\cot^2(\theta/2)) \equiv \frac{|a_1|^2}{\sin^2(\theta/2)} \tag{12.979}$$

yielding $a_1 = N\sin(\theta/2)$. Substituting a_1 into (12.978), we obtain

$$|\,u_A\rangle_- = N\begin{pmatrix}\sin(\theta/2)\\ -\cos(\theta/2)\,e^{i\phi}\end{pmatrix} \quad \text{or} \quad |\,u_A\rangle_- = N\begin{pmatrix}-\sin(\theta/2)\,e^{-i\phi}\\ \cos(\theta/2)\end{pmatrix}. \tag{12.980}$$

In what follows, we are going to use the left-side expressions of (12.976) and (12.980), respectively, as our right-handed and left-handed spinors:

$$|\,u_A\rangle_+ = N\begin{pmatrix}\cos(\theta/2)\\ \sin(\theta/2)\,e^{i\phi}\end{pmatrix}, \qquad |\,u_A\rangle_- = N\begin{pmatrix}\sin(\theta/2)\\ -\cos(\theta/2)\,e^{i\phi}\end{pmatrix}. \tag{12.981}$$

Note that, using (12.969) and (12.981), we can verify that $\widehat{\mathcal{H}}\,|u_A\rangle_\pm = \pm\frac{\hbar}{2}|u_A\rangle_\pm$ (see Problem 12.36 on Page 887).

Now, to find the two-component spinor $|\,u_B\rangle$, we would need to use the Dirac equation for positive energies $(\vec{\alpha}\cdot\vec{p} + mc^2\beta)\,|\,u\rangle = E\,|\,u\rangle$ which, as shown above, can be written as:

$$\begin{pmatrix}mc^2 I_2 & c\vec{\sigma}\cdot\hat{\vec{p}}\\ c\vec{\sigma}\cdot\hat{\vec{p}} & -mc^2 I_2\end{pmatrix}\begin{pmatrix}|\,u_A\rangle\\ |\,u_B\rangle\end{pmatrix} = E\begin{pmatrix}|\,u_A\rangle\\ |\,u_B\rangle\end{pmatrix} \Longrightarrow \begin{cases}(E - mc^2)\,|\,u_A\rangle = cp\vec{\sigma}\cdot\hat{n}\,|\,u_B\rangle\\ cp\vec{\sigma}\cdot\hat{n}\,|\,u_A\rangle = (E + mc^2)\,|\,u_B\rangle\end{cases}, \tag{12.982}$$

where we have used $\hat{\vec{p}} = p\hat{n}$ and where $E = \sqrt{\vec{p}^{\,2}c^2 + m^2c^4}$ (notice that, instead of the notation $E^{(+)}$ utilized above to denote the positive energy, we are shifting here and below to the lighter notation E). Since $\vec{\sigma}\cdot\hat{n}\,|\,u_A\rangle_\pm = \pm\,|\,u_A\rangle_\pm$, the bottom equation in (12.982) allows us to express $|\,u_B\rangle$ in terms of $|\,u_A\rangle$ as follows

$$|\,u_B\rangle_\pm = \pm\frac{cp}{E + mc^2}\,|\,u_A\rangle_\pm. \tag{12.983}$$

Substituting (12.981) into (12.983), we obtain

$$|\,u_B\rangle_+ = N\begin{pmatrix}\frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right)\\ \frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right)e^{i\phi}\end{pmatrix}, \qquad |\,u_B\rangle_- = N\begin{pmatrix}-\frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right)\\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right)e^{i\phi}\end{pmatrix}. \tag{12.984}$$

Finally, to find the eigenstates of the helicity operator (12.965) for the case of particles, we would only need to substitute (12.981) and (12.984) into (12.968):

$$|u\rangle_+ = \begin{pmatrix}|\,u_A\rangle_+\\ |\,u_B\rangle_+\end{pmatrix} = N\begin{pmatrix}\cos\left(\frac{\theta}{2}\right)\\ \sin\left(\frac{\theta}{2}\right)e^{i\phi}\\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right)\\ \frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right)e^{i\phi}\end{pmatrix}, \qquad |u\rangle_- = \begin{pmatrix}|\,u_A\rangle_-\\ |\,u_B\rangle_-\end{pmatrix} = N\begin{pmatrix}\sin\left(\frac{\theta}{2}\right)\\ -\cos\left(\frac{\theta}{2}\right)e^{i\phi}\\ -\frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right)\\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right)e^{i\phi}\end{pmatrix}. \tag{12.985}$$

Helicity eigenstates for antiparticles Following the same method outlined above for particles, we can easily obtain the helicity eigenstates for antiparticles. Combining (12.965), (12.967), and (12.968), we obtain

$$\widehat{\mathcal{H}}|v\rangle_\pm = \mp\frac{\hbar}{2}|v\rangle_\pm \quad \Longrightarrow \quad \begin{pmatrix}\vec{\sigma}\cdot\hat{n} & 0\\ 0 & \vec{\sigma}\cdot\hat{n}\end{pmatrix}\begin{pmatrix}|\,v_A\rangle_\pm\\ |\,v_B\rangle_\pm\end{pmatrix} = \mp\begin{pmatrix}|\,v_A\rangle_\pm\\ |\,v_B\rangle_\pm\end{pmatrix}. \tag{12.986}$$

Taking $|\,v_B\rangle_\pm = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ where b_1 and b_2 are complex numbers and using (12.964), we have

$$\vec{\sigma}\cdot\hat{n}\,|\,v_B\rangle_\pm = \mp\,|\,v_B\rangle_\pm \implies \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \mp\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \tag{12.987}$$

First, let us consider the positive helicity case: $\vec{\sigma}\cdot\hat{n}\,|\,v_B\rangle_+ = -\,|\,v_B\rangle_+$; hence, (12.987) leads to

$$b_1\cos\theta + b_2 e^{-i\phi}\sin\theta = -b_1 \implies b_1(1+\cos\theta) = -b_2 e^{-i\phi}\sin\theta \implies b_2 = -b_1\cot(\theta/2)\,e^{i\phi}. \tag{12.988}$$

Inserting the values of b_1 and b_2 into $|\,v_A\rangle_+ = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, we have

$$|\,v_B\rangle_+ = b_1\begin{pmatrix} 1 \\ -\cot(\theta/2)\,e^{i\phi} \end{pmatrix} \quad \text{or} \quad |\,v_B\rangle_+ = b_2\begin{pmatrix} -\tan(\theta/2)\,e^{-i\phi} \\ 1 \end{pmatrix}. \tag{12.989}$$

Now, we can obtain the numbers b_1 and b_2 by normalizing $|\,v_B\rangle_+$ to N^2:

$$N^2 = {}_+\langle v_B\,|\,v_B\rangle_+ = |b_1|^2\begin{pmatrix} 1 & -\cot(\theta/2)\,e^{-i\phi} \end{pmatrix}\begin{pmatrix} 1 \\ -\cot(\theta/2)\,e^{i\phi} \end{pmatrix} = |b_1|^2(1+\cot^2(\theta/2)) \equiv \frac{|b_1|^2}{\sin^2(\theta/2)}. \tag{12.990}$$

This yields $b_1 = N\sin(\theta/2)$. Inserting b_1 into (12.989), we obtain

$$|\,v_B\rangle_+ = N\begin{pmatrix} \sin(\theta/2) \\ -\cos(\theta/2)\,e^{i\phi} \end{pmatrix} \quad \text{or} \quad |\,v_B\rangle_+ = N\begin{pmatrix} -\sin(\theta/2)\,e^{-i\phi} \\ \cos(\theta/2) \end{pmatrix}. \tag{12.991}$$

Second, let us consider the negative helicity: $\vec{\sigma}\cdot\hat{n}\,|\,v_B\rangle_- = |\,v_B\rangle_-$; hence, (12.987) leads to

$$b_1\cos\theta + b_2 e^{-i\phi}\sin\theta = b_1 \implies b_1(1-\cos\theta) = b_2 e^{-i\phi}\sin\theta \implies b_2 = b_1\tan(\theta/2)\,e^{i\phi}, \tag{12.992}$$

which, when inserted into $|\,v_B\rangle_- = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, yield

$$|\,v_B\rangle_- = b_1\begin{pmatrix} 1 \\ \tan(\theta/2)\,e^{i\phi} \end{pmatrix} \quad \text{or} \quad |\,v_B\rangle_- = b_2\begin{pmatrix} \cot(\theta/2)\,e^{-i\phi} \\ 1 \end{pmatrix}. \tag{12.993}$$

We now can obtain the numbers b_1 and b_2 by normalizing $|\,v_B\rangle_-$ to N^2:

$$N^2 = {}_-\langle v_B\,|\,v_B\rangle_- = |v_1|^2\begin{pmatrix} 1 & \tan(\theta/2)\,e^{-i\phi} \end{pmatrix}\begin{pmatrix} 1 \\ \tan(\theta/2)\,e^{i\phi} \end{pmatrix} = |b_1|^2(1+\tan^2(\theta/2)) \equiv \frac{|b_1|^2}{\cos^2(\theta/2)} \tag{12.994}$$

yielding $b_1 = N\cos(\theta/2)$. Substituting b_1 into (12.978), we obtain

$$|\,v_B\rangle_- = N\begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2)\,e^{i\phi} \end{pmatrix} \quad \text{or} \quad |\,v_B\rangle_- = N\begin{pmatrix} -\sin(\theta/2)\,e^{-i\phi} \\ \cos(\theta/2) \end{pmatrix}. \tag{12.995}$$

In what follows, we are going to use the left-side expressions of (12.991) and (12.995), respectively, as our right-handed and left-handed spinors:

$$|\,v_B\rangle_+ = N\begin{pmatrix} \sin\left(\frac{\theta}{2}\right) \\ -\cos\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}, \qquad |\,v_B\rangle_- = N\begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}. \tag{12.996}$$

Now, to find the two-component spinor $|\, v_A\rangle$, we would need to use the Dirac equation for *negative* energies where we assign $\vec{p} \longrightarrow -\vec{p}$ and $E \longrightarrow -E$ (Feynman-Stückelberg interpretation): $(-\vec{\alpha}\cdot\vec{p} + mc^2\beta)\,|\, u\rangle = -E\,|\, u\rangle$ which, as shown above, can be written as:

$$\begin{pmatrix} mc^2 I_2 & -c\vec{\sigma}\cdot\hat{\vec{p}} \\ -c\vec{\sigma}\cdot\hat{\vec{p}} & -mc^2 I_2 \end{pmatrix}\begin{pmatrix} |\, v_A\rangle \\ |\, v_B\rangle \end{pmatrix} = -E\begin{pmatrix} |\, v_A\rangle \\ |\, v_B\rangle \end{pmatrix} \Longrightarrow \begin{cases} (E + mc^2)\,|\, v_A\rangle = cp\vec{\sigma}\cdot\hat{n}\,|\, v_B\rangle \\ cp\vec{\sigma}\cdot\hat{n}\,|\, v_A\rangle = (E - mc^2)\,|\, v_B\rangle \end{cases}, \tag{12.997}$$

where $E = \sqrt{\vec{p}^{\,2} + m^2c^4}$. The top equation in (12.997) allows us to express $|\, v_A\rangle$ in terms of $|\, v_B\rangle$ as follows

$$|\, v_A\rangle = \frac{c\vec{\sigma}\cdot\hat{\vec{p}}}{E + mc^2}\,|\, v_B\rangle \qquad \Longrightarrow \qquad |\, v_A\rangle_\pm = \mp\frac{cp}{E + mc^2}\,|\, v_B\rangle_\pm, \tag{12.998}$$

where we have used the fact that $\vec{\sigma}\cdot\hat{\vec{p}}\,|\, v_B\rangle_\pm \equiv p(\vec{\sigma}\cdot\hat{n}\,|\, v_B\rangle_\pm) = \mp p\,|\, v_B\rangle_\pm$ in the last equation.

Substituting (12.996) into (12.998), we obtain

$$|\, v_A\rangle_+ = N\begin{pmatrix} -\frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}, \qquad |\, v_A\rangle_- = N\begin{pmatrix} \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}. \tag{12.999}$$

Finally, for the case of antiparticles, the eigenstates of the helicity operator (12.965) can be obtained by substituting (12.996) and (12.999) into $|v\rangle_\pm = \begin{pmatrix} |\, v_B\rangle_\pm \\ |\, v_A\rangle_\pm \end{pmatrix}$:

$$|v\rangle_+ = \begin{pmatrix} |\, v_A\rangle_+ \\ |\, v_B\rangle_+ \end{pmatrix} = N\begin{pmatrix} -\frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right)e^{i\phi} \\ \sin\left(\frac{\theta}{2}\right) \\ -\cos\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}, \qquad |v\rangle_- = \begin{pmatrix} |\, v_A\rangle_- \\ |\, v_B\rangle_- \end{pmatrix} = N\begin{pmatrix} \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right)e^{i\phi} \\ \cos(\theta/2) \\ \sin\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}. \tag{12.1000}$$

In sum, using (12.985) and (12.1000), we can write the particle and anti-particle helicity as well as Hamiltonian eigenstates as follows

$$\boxed{\psi_+(x^\mu) = N\begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right)e^{i\phi} \\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}e^{-i(Et-\vec{p}\cdot\vec{r})}, \quad \psi_-(x^\mu) = N\begin{pmatrix} \sin\left(\frac{\theta}{2}\right) \\ -\cos\left(\frac{\theta}{2}\right)e^{i\phi} \\ -\frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}e^{-i(Et-\vec{p}\cdot\vec{r})}}, \tag{12.1001}$$

$$\boxed{\Phi_+(x^\mu) = N\begin{pmatrix} -\frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right)e^{i\phi} \\ \sin\left(\frac{\theta}{2}\right) \\ -\cos\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}e^{i(Et-\vec{p}\cdot\vec{r})}, \quad \Phi_-(x^\mu) = N\begin{pmatrix} \frac{cp}{E+mc^2}\cos\left(\frac{\theta}{2}\right) \\ \frac{cp}{E+mc^2}\sin\left(\frac{\theta}{2}\right)e^{i\phi} \\ \cos(\theta/2) \\ \sin\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}e^{i(Et-\vec{p}\cdot\vec{r})}}, \tag{12.1002}$$

where $E = \sqrt{\vec{p}^{\,2}c^2 + m^2c^4}$ and the normalization constant N is given by $N = \sqrt{E + mc^2}$.

Remarks

- Notice that the particle and anti-particle helicity eigenstates (12.985) and (12.1000) are also eigenstates to the Dirac Hamiltonian (12.966).

- The particle and anti-particle helicity eigenstates (12.985) and (12.1000) are normalized to $2E$ and orthogonal (see Problem 12.34 on Page 878):

$$ {}_{\pm}\langle u \mid u\rangle_{\pm} = {}_{\pm}\langle v \mid v\rangle_{\pm} = 2E, \qquad {}_{+}\langle u \mid u\rangle_{-} = {}_{+}\langle v \mid v\rangle_{-} = 0. \tag{12.1003} $$

Problem 12.36
Consider a free particle whose momentum is in an arbitrary direction $(\theta,\ \phi)$ that we have studied in Problem 12.35. Using the expressions

$$ \widehat{\mathcal{H}} = \frac{\hbar}{2}\begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta \\ e^{i\phi}\sin\theta & -\cos\theta \end{pmatrix}, \qquad \mid u_A\rangle_{-} = \begin{pmatrix} \sin(\theta/2) \\ -\cos(\theta/2)\, e^{i\phi} \end{pmatrix}, $$

show that the two-component spinor $\mid u_A\rangle_{-}$ is an eigenstate of the helicity operator $\widehat{\mathcal{H}}$ and find the eigenvalue.

Solution
Let us apply the helicity operator on the spinor

$$ \begin{aligned} \widehat{\mathcal{H}} \mid u_A\rangle_{-} &= \frac{\hbar}{2}\begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta \\ e^{i\phi}\sin\theta & -\cos\theta \end{pmatrix}\begin{pmatrix} \sin\left(\frac{\theta}{2}\right) \\ -\cos\left(\frac{\theta}{2}\right) e^{i\phi} \end{pmatrix} \\ &= \frac{\hbar}{2}\begin{pmatrix} \cos\theta\sin\left(\frac{\theta}{2}\right) - \sin\theta\cos\left(\frac{\theta}{2}\right) \\ \left[\sin\theta\sin\left(\frac{\theta}{2}\right) + \cos\theta\cos\left(\frac{\theta}{2}\right)\right] e^{i\phi} \end{pmatrix}. \end{aligned} \tag{12.1004} $$

Now, let us simplify the following two algebraic relations:

$$ \begin{aligned} \cos\theta\sin\left(\frac{\theta}{2}\right) - \sin\theta\cos\left(\frac{\theta}{2}\right) &= \left[\cos^2\left(\frac{\theta}{2}\right) - \sin^2\left(\frac{\theta}{2}\right)\right]\sin\left(\frac{\theta}{2}\right) - \left[2\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)\right]\cos\left(\frac{\theta}{2}\right) \\ &= \cos^2\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) - \sin^3\left(\frac{\theta}{2}\right) - 2\sin\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right) \\ &= -\sin\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right) - \sin^3\left(\frac{\theta}{2}\right) \\ &= -\left[\cos^2\left(\frac{\theta}{2}\right) + \sin^2\left(\frac{\theta}{2}\right)\right]\sin\left(\frac{\theta}{2}\right) \\ &= -\sin\left(\frac{\theta}{2}\right), \end{aligned} \tag{12.1005} $$

and

$$\begin{aligned}\sin\theta\sin\left(\frac{\theta}{2}\right)+\cos\theta\cos\left(\frac{\theta}{2}\right) &= \left[2\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)\right]\sin\left(\frac{\theta}{2}\right)+\left[\cos^2\left(\frac{\theta}{2}\right)-\sin^2\left(\frac{\theta}{2}\right)\right]\cos\left(\frac{\theta}{2}\right)\\ &= 2\sin^2\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)+\cos^3\left(\frac{\theta}{2}\right)-\sin^2\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)\\ &= \sin^2\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)+\cos^3\left(\frac{\theta}{2}\right)\\ &= \cos\left(\frac{\theta}{2}\right) \qquad (12.1006)\end{aligned}$$

Inserting (12.1005) and (12.1006) into (12.1004), we obtain

$$\begin{aligned}\widehat{\mathcal{H}}\ |\,u_A\rangle_- &= \frac{\hbar}{2}\begin{pmatrix}\cos\theta\sin\left(\frac{\theta}{2}\right)-\sin\theta\cos\left(\frac{\theta}{2}\right)\\ \left[\sin\theta\sin\left(\frac{\theta}{2}\right)+\cos\theta\cos\left(\frac{\theta}{2}\right)\right]e^{i\phi}\end{pmatrix}\\ &= \frac{\hbar}{2}\begin{pmatrix}-\sin\left(\frac{\theta}{2}\right)\\ \cos\left(\frac{\theta}{2}\right)e^{i\phi}\end{pmatrix}\\ &= -\frac{\hbar}{2}\,|\,u_A\rangle_- \qquad (12.1007)\end{aligned}$$

Hence, the spinor $|\,u_A\rangle_-$ is an eigenstate of the helicity operator $\widehat{\mathcal{H}}$ with an eigenvalue of $-\hbar/2$.

Problem 12.37
Consider the matrix representations in spherical coordinates of the helicity and Hamiltonian operators that we have derived in Problem 12.35 for the case where the particle's spin is oriented in an arbitrary direction. Show that the Hamiltonian (12.966) and the helicity (12.965) commute.

Solution:

First, the product of the Hamiltonian and helicity matrices is easy to calculate:

$$\begin{aligned}\hat{H}\widehat{\mathcal{H}} &= \frac{cp\hbar}{2}\begin{pmatrix}mc/p & 0 & \cos\theta & e^{-i\phi}\sin\theta\\ 0 & mc/p & e^{i\phi}\sin\theta & -\cos\theta\\ \cos\theta & e^{-i\phi}\sin\theta & -mc/p & 0\\ e^{i\phi}\sin\theta & -\cos\theta & 0 & -mc/p\end{pmatrix}\begin{pmatrix}\cos\theta & e^{-i\phi}\sin\theta & 0 & 0\\ e^{i\phi}\sin\theta & -\cos\theta & 0 & 0\\ 0 & 0 & \cos\theta & e^{-i\phi}\sin\theta\\ 0 & 0 & e^{i\phi}\sin\theta & -\cos\theta\end{pmatrix}\\ &= \frac{mc^2\hbar}{2}\begin{pmatrix}\cos\theta & e^{-i\phi}\sin\theta & 1 & 0\\ e^{i\phi}\sin\theta & -\cos\theta & 0 & 1\\ 1 & 0 & -\cos\theta & -e^{-i\phi}\sin\theta\\ 0 & 1 & -e^{i\phi}\sin\theta & \cos\theta\end{pmatrix} \qquad (12.1008)\end{aligned}$$

Second, we can also easily obtain the helicity and Hamiltonian matrices:

$$\widehat{\mathcal{H}}\,\hat{H} = \frac{cp\hbar}{2}\begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta & 0 & 0 \\ e^{i\phi}\sin\theta & -\cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & e^{-i\phi}\sin\theta \\ 0 & 0 & e^{i\phi}\sin\theta & -\cos\theta \end{pmatrix}\begin{pmatrix} mc/p & 0 & \cos\theta & e^{-i\phi}\sin\theta \\ 0 & mc/p & e^{i\phi}\sin\theta & -cos\theta \\ \cos\theta & e^{-i\phi}\sin\theta & -mc/p & 0 \\ e^{i\phi}\sin\theta & -\cos\theta & 0 & -mc/p \end{pmatrix}$$

$$= \frac{mc^2\hbar}{2}\begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta & 1 & 0 \\ e^{i\phi}\sin\theta & -\cos\theta & 0 & 1 \\ 1 & 0 & -\cos\theta & -e^{-i\phi}\sin\theta \\ 0 & 1 & -e^{i\phi}\sin\theta & \cos\theta \end{pmatrix} \tag{12.1009}$$

A comparison of (12.1008) and (12.1009) shows that $\hat{H}\,\widehat{\mathcal{H}} = \widehat{\mathcal{H}}\,\hat{H}$; hence, the helicity and Hamiltonian operators commute:

$$\left[\hat{H}, \widehat{\mathcal{H}}\right] = 0. \tag{12.1010}$$

12.8 Exercises

Exercise 12.1
(a)Using the classical formula of kinetic energy, estimate the amount of energy needed to accelerate the 1 kg object to a relativistic speed of 2.9×10^8 m/s. Give an idea on the size of this energy in terms of the yield of a thermonuclear bomb. Note: The yield of a thermonuclear device is measured in terms of megatons of the conventional explosive, TNT; the conversion formula is: 1 Megaton of TNT $\simeq 4.184\times10^{15}J$ (Joules).
(b) Calculate the rest-mass energy of a $m_0 = 1.00$ kg object. Again, give an idea on the size of this energy in terms of the yield of a thermonuclear device.

Exercise 12.2
(a) Find the Minkowski metric for a 4-dimensional spacetime in polar coordinates.
(b) Find the inverse of the metric derived in (a).

Exercise 12.3
Using the Minkowski metric (12.18), show that $g_{\mu\nu}g^{\mu\nu} = 4$.

Exercise 12.4
Using the Dirac γ−matrices, show that the square root of the negative of the d'Alembertian operator is given by

$$\left(i\gamma^\mu\partial_\mu\right)^2 = -\Box. \tag{12.1011}$$

Exercise 12.5
Consider two inertial frames of reference F and F' where F' is in uniform motion along the y−axis with constant velocity $\vec{v}$ relative to F. An event is described by the 4-coordinates $(t,\ x,\ y,z)$ in F and by $(t',\ x',\ y',\ z')$ in F'. Find the Galilean transformation that connects $(t,\ x,\ y,z)$ to $(t',\ x',\ y',\ z')$.

Exercise 12.6
Consider two inertial frames of reference F and F' where F' is in uniform motion along the $y-$axis with constant velocity $\vec{v}$ relative to F. An event is described by the 4-coordinates $(t,\ x,\ y,z)$ in F and by $(t',\ x',\ y',\ z')$ in F'. Find the *inverse* Galilean transformation; that is, the transformation that connects $(t',\ x',\ y',\ z')$ to $(t,\ x,\ y,z)$.

Exercise 12.7
Consider two inertial frames of reference F and F' where F' is in uniform motion along the $y-$axis with constant velocity $\vec{v}$ relative to F. An event is described by the 4-coordinates $(ct,\ x,\ y,z)$ in F and by $(ct',\ x',\ y',\ z')$ in F'. Find the *inverse* Lorentz transformation that connects $(ct',\ x',\ y',\ z')$ to $(ct,\ x,\ y,z)$.

Exercise 12.8
Consider two inertial frames of reference F and F' where F' is in uniform motion along the $z-$axis with constant velocity $\vec{v}$ relative to F. An event is described by the 4-coordinates $(ct,\ x,\ y,z)$ in F and by $(ct',\ x',\ y',\ z')$ in F'. Find the Lorentz transformation that connects $(ct,\ x,\ y,z)$ to $(ct',\ x',\ y',\ z')$.

Exercise 12.9
Consider two inertial frames of reference F and F' where F' is in uniform motion along the $y-$axis with constant velocity $\vec{v}$ relative to F. Two events are described by the 4-vectors $A^\mu = (A^0,\ A^1,\ A^2,\ A^3)$ and $B^\mu = (B^0,\ B^1,\ B^2,\ B^3)$ in F and by $A'^\mu = (A'^0,\ A'^1,\ A'^2,\ A'^3)$ and $B'^\mu = (B'^0,\ B'^1,\ B'^2,\ B'^3)$ in F'. Using the boost (Lorentz transformation) along the $y-$axis

$$\Lambda = \begin{pmatrix} \cosh\phi & 0 & -\sinh\phi & 0 \\ 0 & 1 & 0 & 0 \\ -\sinh\phi & 0 & \cosh\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where $\tanh\phi = v/c$ is the rapidity, show that the inner product between any two 4-vectors, $A^\mu = (A^0,\ A^1,\ A^2,\ A^3)$ and $B^\mu = (B^0,\ B^1,\ B^2,\ B^3)$, is invariant under this boost; that is, show that:

$$A'^\mu B'_\mu = A^\mu B_\mu.$$

Exercise 12.10
Consider the following two consecutive boosts: a v_y boost along the $y-$direction followed by a v_z boost along the $z-$direction.

(a) Find the Lorentz transformations corresponding to the two boosts described above.

(b) Show that when the two boosts described above are performed in the reverse order, they would give a different transformation; that is, show that the Lorentz transformation corresponding to a v_z boost in the $z-$direction followed by a v_y boost in the $y-$direction would be different from the result obtained in (a).

Exercise 12.11
Consider two inertial frames F and F' where F' is in uniform motion along the y-direction with constant speed $v = 2.4\times10^8$ m/s relative to F. Consider also two events that take place in F and

whose spacetime coordinates are given by the following 4-vectors in Cartesian coordinates in the Minkowski space time: $A^\mu = (7,3,0,2)$, $B^\mu = (3,1,1,-2)$.

(a) Find the matrix of the Lorentz transformation corresponding to this $y-$axis boost.

(b) Find the coordinates A'^μ and B'^μ of these two events as measured by an observer in F'.

(c) Show quantitatively that the squares of the 4-vectors A^μ and B^μ are Lorentz invariant.

(d) Show quantitatively that the inner product between A^μ and B_μ is a Lorentz invariant.

Exercise 12.12

Consider two inertial frames F and F' where F' is in uniform motion along the z-direction with constant speed $v = 2.8 \times 10^8$ m/s relative to F. Consider also two events that take place in F and whose Minkowski spacetime coordinates are given by the following 4-vectors: $A^\mu = (4,-1,3,2)$, $B^\mu = (6,-3,1,-4)$.

(a) Find the matrix of the Lorentz transformation corresponding to this $y-$axis boost.

(b) Find the coordinates A'^μ and B'^μ of these two events as measured by an observer in F'.

(c) Show quantitatively that the squares of the 4-vectors A^μ and B^μ are Lorentz invariant.

(d) Show quantitatively that the inner product between A^μ and B_μ is a Lorentz invariant.

Exercise 12.13

Consider the following 4-vectors in Minkowski spacetime with components in Cartesian coordinates as given by: $A^\mu = (3,-1,2,1)$, $B^\mu = (2,3,2,-1)$, $C^\mu = (-2,2,-2,0)$.

(a) Determine the causal nature (timelike, spacelike, or null) of A^μ, B^μ, and C^μ.

(b) Are any two of these vectors orthogonal to each other?

Exercise 12.14

Are any two of the following 4-vectors orthogonal to each other in the Minkowski spacetime?

$$A^\mu = (-1,0,1,2), \qquad B^\mu = (0,-2,1,0), \qquad C^\mu = (1,0,0,-3).$$

Exercise 12.15

Determine the causal nature (timelike, spacelike, or null) of the following 4-vectors

$$A^\mu = \left(0, \frac{1}{\sqrt{2}}, 1, -\frac{1}{\sqrt{2}}\right), \quad B^\mu = \left(\frac{1}{2}, -2, 1, \frac{1}{2}\right), \quad C^\mu = (3,1,0,-1), \quad D^\mu = (-4, \sqrt{3}, -2, 3).$$

Exercise 12.16

Consider two inertial frames F and F' where F' is in uniform motion along the y-direction with constant speed $v = 2.6 \times 10^8$ m/s relative to F. Consider also two events that take place in F and whose Minkowski spacetime coordinates are given by the following 4-vectors: $A^\mu = (3,2,-4,1)$, $B^\mu = (-2,-1,3,2)$.

(a) Find the matrix of the Lorentz transformation corresponding to this $x-$axis boost.

(b) Find the coordinates A'^μ and B'^μ of these two events as measured by an observer in F'.

(c) Show that this Lorentz transformation preserves the causal nature (timelike, spacelike, or null) of A^μ and B^μ.

(d) Show quantitatively that the distance between these two events is a Lorentz invariant.

Exercise 12.17

Consider a wave function $\psi(x^\mu) = a^\mu \exp\left(-i(\omega t - \vec{k}\cdot\vec{r})\right)$, where a^μ is a constant 4-vector, ω is the angular frequency, and $\vec{k}$ is the wave vector.

(a) Show that, if $\omega = c|\vec{k}|$, then $\Box\psi(x^\mu) = 0$, where $\Box$ is the d'Alembert operator.

(b) Verify that $\psi(x^\mu)$ is an eigenstate of the operators $\hat{H} = i\hbar\partial/\partial t$ and $\hat{\vec{p}} = -i\hbar\vec{\nabla}$ with eigenvalues $\hbar\omega$ and $\hbar\vec{k}$, respectively.

Exercise 12.18
Find the solutions of the Klein–Gordon equation for a free particle at rest.

Exercise 12.19
Find the solutions of the Klein–Gordon equation for a free particle moving in the xy–plane with non-zero momentum, $\vec{p} = p_x\hat{i} + p_y\hat{j}$.

Exercise 12.20
Using the matrix representations of the γ^μ matrices listed in Eqs. (12.222)–(12.223), verify that the relation $\gamma^\nu\gamma^\mu\gamma_\nu = -2\gamma^\mu$ holds for the case $\mu = 3$.

Exercise 12.21
Using the matrices of γ^μ listed in Eqs. (12.222)–(12.223), verify that the relation $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4$ holds for the case $\mu = 1$ and $\nu = 3$ where $[g]$is the Minkowski metric; i.e., $g^{00} = 1$, $g^{11} = g^{22} = g^{33} = -1$, and $g^{\mu\nu} = 0$ when $\mu \neq \nu$.

Exercise 12.22
Consider the following three 4-vectors: $a^\mu = (-5, 3, -3, 1)$, $b^\mu = (2, -1, 1, 0)$, and $c^\mu = (4, -3, 3, 2)$.

(a) Find the slashed 4-vectors: $\not{a}, \not{b}, \not{c}$.

(b) Show that Tr $(\not{a}) = 0$.

(c) Calculate Tr $\left(\gamma^5\not{b}\right)$.

(d) Calculate Tr $\left(\not{a}\not{b}\not{c}\right)$.

Exercise 12.23
Using the matrices of γ^μ listed in text, verify that the matrix of $\gamma^5 = i\gamma_0\gamma_1\gamma_2\gamma_3$ is given by

$$\gamma^5 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Exercise 12.24
Using the definition $\gamma^5 = i\gamma_0\gamma_1\gamma_2\gamma_3$, verify that the relation $(\gamma^5)^2 = I_4$ holds by pursuing the following two approaches:

(a) First, by using an algebraic approach involving the properties of the γ^μ matrices.

(b) Second, by using the matrices of γ^μ listed in Eqs. (12.222) – (12.224)).

Exercise 12.25
Verify the relation $\gamma^\mu\gamma_\mu = 4I_4$, where I_4 is the four-rowed identity matrix, by using the following two methods:

(a) First, by using an algebraic approach involving the properties of the γ^μ matrices.

(b) Second, by using the matrices of γ^μ listed in Eqs. (12.222) – (12.224).

Exercise 12.26
Using the matrices of γ^μ listed above in equations (12.222) – (12.223) verify that

(a) the relation $(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0$ (listed in (12.176)) holds for $\mu = 3$; i.e., verify that $\gamma^0\gamma^3\gamma^0$ is Hermitian, and that

(b) $\text{Tr}\left(\gamma^0\gamma^3\gamma^0\right) = 0$

Exercise 12.27
Using the matrices of γ^μ and γ^5 listed the main text, calculate the following Poisson brackets

(a) $\left\{\gamma^0,\ \gamma^5\right\}$, and

(b) $\left\{\gamma^2,\ \gamma^5\right\}$.

Exercise 12.28
Using the matrices of γ^0, γ^1, γ^2, γ^3 and γ^5 listed in the main text, calculate the following traces

(a) $\text{Tr}\left(\gamma^0\gamma^1\gamma^2\right)$,

(b) $\text{Tr}\left(\gamma^3\gamma^5\right)$, and

(c) $\text{Tr}\left(\gamma^0\gamma^1\gamma^2\gamma^3\gamma^5\right)$.

Exercise 12.29
Using the matrices of γ^μ and γ^5 listed above in equations (12.222) – (12.224), verify that the trace $\text{Tr}\left(\gamma^\mu\gamma^5\right)$ is zero for the following two cases:

(a) $\mu = 0$, and

(b) $\mu = 2$.

Exercise 12.30
Without using the explicit matrix representation of the gamma matrices, show that $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ satisfies the following properties:

(a) $\gamma^5\gamma^\mu = -\gamma^\mu\gamma^5$, and

(b) $\left(\gamma^5\right)^\dagger = \gamma^5$

Exercise 12.31
Consider the 4×4 matrix $U = \frac{1}{\sqrt{2}}\begin{pmatrix} I_2 & I_2 \\ I_2 & -I_2 \end{pmatrix}$, where I_2 is the 2×2 unity matrix.

(a) Show the U is unitary.

(b) Using the Weyl basis matrices (12.226) and the Dirac basis matrices (12.222) – (12.223), show that

$$\gamma^0_W = U\gamma^0_D U^{-1},$$

where γ^0_D and γ^0_M are the Dirac and the Weyl matrices, respectively.

Exercise 12.32

Using the 4×4 matrix $U = \begin{pmatrix} I_2 & \sigma_y \\ -\sigma_y & I_2 \end{pmatrix}$, where I_2 is the 2×2 unity matrix and σ_y is the y–component of the Pauli matrix, along with the Majorana basis matrices and the Dirac basis matrices listed in the main text, show that

$$\gamma^1_M = U\gamma^1_D U^{-1},$$

where γ^1_D and γ^1_M are the first components of the Dirac matrix and the Majorana matrix, respectively.

Exercise 12.33
Find the solutions of the Dirac equation – energies and eigenstates – for a (relativistic) free particle moving in the xy-plane with a non-zero momentum, $\vec{p} = p_x\hat{i} + p_y\hat{j}$.

Exercise 12.34
Find the solutions of the Dirac equation – energies and eigenstates – for a (relativistic) free particle moving in the z-direction with a non-zero momentum, $\vec{p} = p\hat{k}$.

Exercise 12.35
(a) Verify that the Dirac spinors u_1, u_2, v_1, v_2 listed in Eq. (12.336),

$$u_1 = N\begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E^{(-)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \end{pmatrix}, \quad u_2 = N\begin{pmatrix} 0 \\ 1 \\ \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \end{pmatrix}, \quad v_1 = N\begin{pmatrix} \frac{c(p_x-ip_y)}{E^{(+)}+mc^2} \\ \frac{-cp_z}{E^{(+)}+mc^2} \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = N\begin{pmatrix} \frac{cp_z}{E^{(+)}+mc^2} \\ \frac{c(p_x+ip_y)}{E^{(+)}+mc^2} \\ 1 \\ 0 \end{pmatrix}, \tag{12.1012}$$

are normalized to $2E^{(+)}$, provided $N = \sqrt{E^{(+)} + mc^2}$; i.e., verify that $\langle u_n|u_n\rangle = 2E^{(+)}$ and $\langle v_n|v_n\rangle = 2E^{(+)}$ where $n = 1, 2$.
(b) Verify that the particle spinors u_1 and u_2 are orthogonal and so are the antiparticle spinors v_1 and v_2.

Exercise 12.36
Show that the Schrödinger equation is covariant under time reversal and find the time reversed wave function.

Exercise 12.37
Show that the Klein–Gordon equation is covariant under time reversal and find the time reversed wave function.

Exercise 12.38
Show that the Klein–Gordon equation is covariant under parity.

Exercise 12.39
Show that the Schrödinger equation is covariant under parity.

Exercise 12.40
Using the properties of the γ–matrices and the relations $\hat{P}_R = \frac{1}{2}(I_4 + \gamma^5)$ and $\hat{P}_L = \frac{1}{2}(I_4 - \gamma^5)$, show that $\hat{P}_R$ and $\hat{P}_L$ are projection operators.

Exercise 12.41
Using expressions of the helicity operator, $\widehat{\mathcal{H}}$, and the two-component spinor, $|\, u_A\rangle_+$, in spherical coordinates,

$$\widehat{\mathcal{H}} = \frac{\hbar}{2}\begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta & 0 & 0 \\ e^{i\phi}\sin\theta & -\cos\theta & & \end{pmatrix}, \qquad |\, u_A\rangle_+ = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right)e^{i\phi} \end{pmatrix}$$

show that $\widehat{\mathcal{H}}|u_A\rangle_+ = \frac{\hbar}{2}|u_A\rangle_+$.

Exercise 12.42
Consider a spinless (relativistic) particle with an energy E traveling along the positive z–axis and striking a barrier potential of height V_0,

$$V(z) = \begin{cases} 0, & z \leq 0, \\ V_0, & 0 < z < a, \\ 0, & z \geq a, \end{cases} \tag{12.1013}$$

such that $-mc^2 < V_0 < mc^2$, where mc^2 is the rest-mass energy of the particle.

Find the reflection and transmission coefficients for the case where the particle's energy is in the range of $V_0 - mc^2 < E < V_0$. Also find the transmitted probability and examine if it is negative or positive?

Exercise 12.43
Consider a spin-up (relativistic) free electron with an energy E traveling along the positive z–axis and striking a barrier potential of height V_0,

$$V(z) = \begin{cases} 0, & z \leq 0, \\ V_0, & 0 < z < a, \\ 0, & z \geq a, \end{cases} \tag{12.1014}$$

such that $-mc^2 < V_0 < mc^2$, where mc^2 is the rest-mass energy of the particle.

Discuss the the solutions of the Dirac equation for the following energy ranges:
(a) $E > V_0 + mc^2$,
(b) $E < V_0 - mc^2$.

Chapter 13

Beyond Relativistic Quantum Mechanics

13.1 Introduction

As we saw in the previous chapter, inspired by the success of the Schrödinger equation in the non-relativistic domain, Klein-Gordon and Dirac had one aim: to come up with a *single-particle* equation that will be the basis of relativistic quantum mechanics. Neither the Klein-Gordon nor the Dirac equations has managed to achieve that aim. Following in the footsteps of Schrödinger, Klein-Gordon and Dirac have started with *single-particle* wave functions in their attempts to describe the dynamics of a relativistic particle of mass m; paradoxically, they unwittingly ended up dealing with multiparticle theories, notwithstanding the fact that neither of their equations had the proper theoretical tools to handle multiparticle systems. On the one hand, the problems that have plagued the Klein-Gordon equation were not overcome until it was re-interpreted as a *field equation* for spinless particles; it was shown to work for multi-particle systems and that its solutions include *scalar fields* whose quanta are spin-integer particles. On the other hand, to explain the negative energy solutions of his equation, Dirac had to postulate the existence of an infinite continuum of negative energy states. Additionally, and as we saw in Chapter 12 (Subsections 12.3.5 and 12.4.12), the Klein paradox encountered in both equations has shown clearly that a *single-particle* quantum mechanical description of relativistic phenomena is simply inadequate; both of the two equations ended up dealing with the creation of particle-antiparticle pairs.

To address the limitations of relativistic quantum mechanics, one has to go beyond the single-particle character of the Klein-Gordon and the Dirac equations and invoke a theory that must be *multiparticle* in nature. The answer is provided by *Quantum field theory* (QFT). This theory has been developed from a *combination of classical field theory, special relativity, and quantum mechanics*. QFT, which is much more general than relativistic quantum mechanics, offers the proper framework for dealing with relativistic quantum phenomena and multi-particle systems. Since QFT cannot be covered in a single chapter, we are going to content ourselves with a short outline of *classical field theory* and how it can be used to derive the Klein-Gordon and Dirac equations. But why do we need to study classical field theory? Much like classical mechanics which serves as a pre-requisite for studying quantum mechanics, classical

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

field theory also serves as a prelude to quantum field theory. Classical field theory provides the necessary concepts and theoretical tools to smoothly segway into quantum field theory. In short, classical field theory is merely a warm-up for quantum field theory.

Classical field theory deals with the description of how fields interact with particles (matter) by using field equations. For instance, the gravitational and electromagnetic fields are best described by means of classical field theory. In fact, Newton's theory of gravity and Maxwell's theory of electromagnetism are considered to be the very first classical field theories that were introduced in physics. We should note that classical field theories do not deal with quantization; a field theory that incorporates quantization is called a quantum field theory. So, the transition from classical field theory to quantum field theory is achieved by quantizing the classical fields; this topic is not covered in the text. We should mention, however, that, in spite of not being quantum mechanical, classical field theory offers an elegant and straightforward way to deriving the Klein-Gordon and Dirac equations. This is the main goal of the current chapter.

How to mathematically formulate classical field theory? For instance, and as presented in Appendix F, Lagrangians are known to yield a formulation of classical mechanics that is more powerful and elegant than the standard approach of Newton. Similarly, and as we will see below, Lagrangians have become the method of first choice in the formulation of field theories. This is due to a number of reasons, chiefly among them is the fact that Lagrangians lends themselves quite readily to Lorentz covariance, which is recognized to be a fundamental property of nature. That is why modern formulations of classical field theories are constructed around the requirement of achieving Lorentz covariance and making use of Lagrangians. A Lagrangian formulation of classical field theory is known as a *Lagrangian field theory*; this is the theoretical equivalent of Lagrangian mechanics (see Appendix F for a brief outline of Lagrangian mechanics). While Lagrangian mechanics applies to systems of material particles involving finite degrees of freedom, Lagrangian field theory deals with fields or continuous media with infinite degrees of freedom.

After discussing the advantages of the Lagrangian formulation of field theories and offering a brief outline of classical field theory, we will show how to use Lagrangians to derive the Klein-Gordon and Dirac equations.

13.2 Advantages of Lagrangians in Field Theories

As we will see below, in deriving the Klein-Gordon and Dirac equations, we will start from a variational principle and then apply it to the Klein-Gordon and Dirac Lagrangian densities; variational principles are the bedrock upon which fundamental theories are built. The Lagrangian (or field) derivation of the Klein-Gordon and Dirac equations offers a number of theoretical advantages over the traditional derivations that we saw in Chapter 12. Being multi-particle in nature, the field or Lagrangian formulations of the Klein-Gordon and Dirac equations enable us to avoid most of the conceptual as well as the technical difficulties encountered in the traditional, single-particle approaches that we have covered in the previous chapter.

The Lagrangian approach to field theory has a number of inherent advantages. For instance, the Lagrangian is a scalar function, since, as we will see below, the Euler-Lagrange equations involve only scalar quantities; thus, the Euler–Lagrange equations are invariant under Lorentz transformations. Additionally, the Lagrangian formulation handles symmetries

Table 13.1: Lagrangian symmetries and their corresponding conserved quantities.

Lagrangian symmetry	Conserved quantity
Invariance under space translation	Linear momentum
Invariance under space rotations	Angular momentum
Invariance under time translation	Total energy

with ease and clarity: it provides an effective tool to connect symmetries with conservation laws. If the Lagrangian is invariant under a symmetry, the resulting equations of motion are also invariant under that particular symmetry. The underlying power of the Lagrangian formalism resides in the one to one correspondence between Lagrangian symmetries and constants of the motion. This appealing feature of the Lagrangian formalism is a consequence of the *Noether theorem* which indicates that *every symmetry of the Lagrangian has a corresponding conservation law*; there is a one to one correspondence between the various Lagrangian symmetries and their corresponding conserved quantities. As illustrated in Table 13.1, the invariance of the Lagrangian under spatial translations, spatial rotations, and time translations lead, respectively, to the conservation of linear momentum, angular momentum, and energy. Another noteworthy advantage of the Lagrangian formulation is the ease with which it can be extended to relativistic systems. One final advantage of the Lagrangian formalism is the fact that it lends itself quite readily to the quantization process, for it leads to the Hamiltonian formulation which, in turn, allows for a smooth transition to quantum mechanics through the well known procedure of canonical quantization as illustrated in Appendix F (see Subsection F.3).

Before dealing with the Lagrangian derivations of the Klein-Gordon and Dirac equations, we are going to offer first a brief outline of classical field theory; in this context, the reader may find Appendix F useful to acquire the necessary tools to study classical field theory.

13.3 Lagrangians in Classical Field Theory

In addition to offering an accurate description of classical mechanics from first principles, Lagrangian mechanics applies to field theory as well. The role played by Lagrangians in field theory is as important as that of Lagrangians in classical particle mechanics (see Appendix F). Although the Lagrangian formulation is not the only method to formulate a field theory, they have become the theoretical tool of choice for dealing with field theories, for they lend themselves quite readily to the Lagrangian formulation. One of the quintessential reasons is that the Lagrangian formulation makes it easy to see whether a field theory is invariant or not under various symmetries. If the Lagrangian is Lorentz invariant, the field theory will forcibly be Lorentz invariant as well. Furthermore, using *Noether's theorem* (which states that every continuous symmetry of the Lagrangian has a corresponding conservation law), one would be able to derive important conservation laws of a field theory from the symmetries of the Lagrangian.

Unlike classical mechanics which deals with *discrete* sets of particles and, hence, *finite numbers of degrees of freedom*, classical field theory deals with *continuous* entities — *fields* — that involve *infinite numbers of degrees of freedom*; it treats particles as excited states of their corresponding fields. While a particle is localized in space, a field occupies or extends over

a region in space-time. So, a field $\phi(x^\mu)$ is a quantity defined at every point of space-time $x^\mu = (ct, \vec{r})$. The electric and magnetic fields, $\vec{E}(t, \vec{r})$ and $\vec{B}(t, \vec{r})$, are the most familiar examples of classical fields; as seen in Appendix E, they can be combined into a single 4-component field $A^\mu = (A^0, \vec{A})$ known as the electromagnetic field. The 4-vector field A^μ describes the coupling of a particle of charge q (and electric potential U) to an electromagnetic field in terms of the 3-dimensional vector potential $\vec{A}$: $A^\mu = (A^0, \vec{A}) = \left(U/c, \vec{A}\right)$.

In what follows, we are going to offer a brief review of the most important features of classical field theory, most notably the field Lagrangian formulation of classical field theory that will be relevant to the derivation of the Klein-Gordon and Dirac field equations.

By analogy to classical mechanics, we are going to use Lagrangians to describe the dynamics of a field. While the Lagrangian of a system is expressed in classical mechanics in terms of a *discrete* and *finite* set of generalized coordinates, q_j, and their time derives, $\dot{q}_j$, the Lagrangian is expressed in field theory in terms of *continuous functions*. The transition from classical mechanics to classical field theory is schematically achieved by replacing q_j and $\dot{q}_j$ with $\phi(x^\mu)$ and $\partial_\mu\phi(x^\mu) \equiv \frac{\partial\phi(x^\mu)}{\partial x^\mu}$, respectively:

$$q_j \longrightarrow \phi(x^\mu), \qquad \dot{q}_j \longrightarrow \partial_\mu\phi. \tag{13.1}$$

So, in field theory, the Lagrangian depends at a space-time point $x^\mu = (ct, x, y, z)$ on the values of the fields $\phi(x^\mu)$ and their space-time derivatives $\partial_\mu\phi(x^\mu)$, but not explicitly on x^μ:

$$L = L(\phi, \partial_\mu\phi). \tag{13.2}$$

We can write the Lagrangian (13.2) as a volume integral over the 3-dimensional space as follows

$$L(\phi, \partial_\mu\phi) = \int d^3x\, \mathcal{L}(\phi, \partial_\mu\phi), \tag{13.3}$$

where $d^3x = dxdydz$ and where $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu\phi)$ is the *Lagrangian density*, which depends at a point x^μ only on the values of the fields $\phi(x^\mu)$ and their derivatives $\partial_\mu\phi(x^\mu)$ at this point; thus, $\mathcal{L}$ must be local and the theory is called a local field theory. Since the Lagrangian density does not depend explicitly[1] on x^μ, it is invariant under space-time translations and, hence, linear momentum and energy are conserved. In classical mechanics, we deal with Lagrangians, $L(q, \dot{q})$, whereas in field theory, we deal with Lagrangian densities, $\mathcal{L}(\phi, \partial_\mu\phi)$. By analogy to classical mechanics, the Lagrangian density in field theory can be constructed in terms of a kinetic energy density minus a potential energy density; the Klein-Gordon and Dirac Lagrangian densities that we will see below (see Eqs. (13.63) and (13.88)) provide illustrations.

To obtain the equation of motion of a system, we would need to minimize the following integral, known as the *action* the system, between two endpoints[2] $x_1 = (ct_1, x_1, y_1, z_1)$ and $x_2 = (ct_2, x_2, y_2, z_2)$ that are fixed:

$$S = \int_{t_1}^{t_2} dt\, L(\phi, \partial_\mu\phi; t) = \int_{t_1}^{t_2} dt \int d^3x\, \mathcal{L}(\phi, \partial_\mu\phi) \equiv \frac{1}{c}\int_{x_1}^{x_2} d^4x\, \mathcal{L}(\phi, \partial_\mu\phi), \tag{13.4}$$

[1] For closed systems, the Lagrangian density does not depend explicitly on the space-time coordinates x^μ, leading to conservation relations in the space and time degrees of freedom; i.e., conservation of linear momentum and energy.

[2] In a relativistic theory such as relativistic quantum field theory, one must treat time and space coordinates on an equal footing.

where $d^4x = dx^0d^3x = cdtd^3x \equiv cdtdxdydz$ is the 4-volume element of the 4-dimensional space-time. Our main task is to find the field ϕ such that the action S is an extremum (minimum or maximum); we will come back to this issue shortly.

A small variation, $\delta\phi$, of the field, ϕ,

$$\phi(x^\mu) \longrightarrow \phi(x^\mu) + \delta\phi(x^\mu), \tag{13.5}$$

leads to a variation of the Lagrangian density, $\mathcal{L} \longrightarrow \mathcal{L} + \delta\mathcal{L}$, where

$$\delta\mathcal{L}(\phi, \partial_\mu\phi) = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi), \tag{13.6}$$

and a variation of the action, $S \longrightarrow S + \delta S$, along an *arbitrary* path between the endpoints x_1 and x_2, where

$$\delta S = \frac{1}{c}\int_{x_1}^{x_2} d^4x\, \delta\mathcal{L} = \frac{1}{c}\int d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi)\right]. \tag{13.7}$$

Although we have already introduced the symbol "δ", which represents a "variation," we want to highlight two of its most important properties: (a) it is linear, and (b) it commutes with the differentiation operation:

$$\delta(\partial_\mu\phi) = \partial_\mu(\delta\phi). \tag{13.8}$$

So, using the fact that δ and ∂_μ commute along with the relation

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi), \tag{13.9}$$

and integrating by parts the last term of (13.7), we can write

$$\int_{x_1}^{x_2} d^4x\, \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi) \equiv \int_{x_1}^{x_2} d^4x\, \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\Big|_{x_1}^{x_2} - \int d^4x\, \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\delta\phi. \tag{13.10}$$

Inserting (13.10) into (13.7), we obtain

$$\delta S = \frac{1}{c}\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\Big|_{x_1}^{x_2} + \frac{1}{c}\int d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right]\delta\phi. \tag{13.11}$$

As mentioned above, since we are varying the action along an arbitrary path between two fixed endpoints x_1 and x_2 but not at the endpoints x_1 and x_2, the variation $\delta\phi(x)$ must vanish at the endpoints, $\delta\phi(x_1) = \delta\phi(x_2) = 0$; that is, the boundary conditions dictate that

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\Big|_{x_1}^{x_2} = 0, \tag{13.12}$$

and, hence, Equation (13.11) becomes

$$\delta S = \frac{1}{c}\int d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right]\delta\phi. \tag{13.13}$$

For the action to be *stationary*[3] (i.e., $\delta S = 0$), the *least (stationary) action principle*, also called *Hamilton's principle of least action*, dictates that, when a system evolves in time from one state to another, the system will follow a path such as *the action is an extremum* (maximum or minimum) for *arbitrary* small variations $\delta\phi$ from the true field ϕ:

$$\delta S = \frac{1}{c}\int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right)\right]\delta\phi = 0. \tag{13.14}$$

Hence, in order for δS to be zero for any small field variation, $\delta\phi$, the following condition must be satisfied:

$$\boxed{\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right) = 0.} \tag{13.15}$$

This is known as the *Euler–Lagrange equation* of motion (i.e., a field equation). This equation is the field analog of the single particle equation (F.7) of classical mechanics that we have derived in Appendix F: $\frac{\partial L}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = 0$, where $j = 1, 2, 3, \ldots$.

Remarks

- By analogy to classical or Lagrangian mechanics where we encounter physical quantities that are scalars, vectors, or tensors, in classical field theory we also encounter physical fields that are *classical, vector, tensor, or spinnor fields*, depending on the system under study. In general, while bosonic systems are described by scalar, vector, or tensor fields, fermions are described by spinnor fields. For instance, we will see in the next two sections how to derive the Klein-Gordon equation from a scalar field and the Dirac equation from a spinnor field.

- As seen above, the Lagrangian for a single field ϕ is represented by $L(\phi, \partial_\mu\phi)$. When dealing with a system involving many fields, $\phi_1, \phi_2, \phi_3, \ldots$, the Lagrangian is represented by $L(\phi_1, \partial\phi_1; \phi_2, \partial\phi_2;, \phi_3, \partial\phi_3; \cdots)$.

In the following three sections, we are going to use the Euler–Lagrange equation (13.15) to obtain field equations for vector fields (electromagnetic fields), real scalar fields (Klein-Gordon fields), and spinor fields (Dirac fields). In order to obtain field equations that are relativistically covariant, the action, S, has to be a Lorentz scalar so that the least action principle will be Lorentz frame independent. However, since the 4-volume element, d^4x, is a (Lorentz) scalar, the Lagrangian density, $\mathcal{L}$, is also a scalar. This requirement will impose constraints on the form of Lagrangian densities that are allowed in field theory. So in the next three sections, we are going to see how to construct Lorentz invariant Lagrangian densities for the electromagnetic, Klein-Gordon, and Dirac fields with the aim of recovering the Maxwell, Klein-Gordon and Dirac field equations.

[3] Another way to look at this is that, within the context of special relativity, we seek field theories where actions are relativistically invariant, since, in accordance with the first postulate of special relativity, the laws of physics are the same in all inertial reference frames. In these theories, the equations of motion are derived from the condition that the action is a Lorentz invariant, $\delta S = 0$.

13.4 Lagrangian Derivation of the Maxwell Equations

An interesting application of the Largrangian formulation of classical field theory is to the derivation of the Maxwell equations. Before looking into this, we are going to give a brief outline of the Maxwell equations in differential form and then in the covariant notation.

13.4.1 Maxwell Equations in Differential Form

From the theory of classical electrodynamics, the electric and magnetic fields, $\vec{E}$ and $\vec{B}$, created by a source (i.e., a charge density, ρ, and a current density, $\vec{J}$) are governed by the *Maxwell equations* and which are given in SI units by:

$$\text{Gauss's law for electricity:} \qquad \vec{\nabla}\cdot\vec{E} = \frac{\rho}{\varepsilon_0}, \tag{13.16}$$

$$\text{Ampere's law:} \qquad \vec{\nabla}\times\vec{B} - \varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t} = \mu_0\vec{J}, \tag{13.17}$$

$$\text{Gauss's law for magnetism:} \qquad \vec{\nabla}\cdot\vec{B} = 0, \tag{13.18}$$

$$\text{Maxwell-Faraday induction law:} \qquad \vec{\nabla}\times\vec{E} + \frac{\partial\vec{B}}{\partial t} = 0. \tag{13.19}$$

Equations (13.16) and (13.17) are commonly referred to as the *inhomogeneous* (or source-generated) Maxwell equations. Note that, since $\varepsilon_0\mu_0 = 1/c^2$, we can also write (13.17) as $\vec{\nabla}\times\vec{B} - \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t} = \mu_0\vec{J}$.

In a source-free region of space (vacuum) with no charges, $\rho = 0$, and no currents, $\vec{J} = 0$, Equations (13.16)–(13.19) reduce to the sourceless or *homogenous* Maxwell equations:

$$\vec{\nabla}\cdot\vec{E} = 0, \qquad \vec{\nabla}\times\vec{B} - \varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t} = 0, \tag{13.20}$$

$$\vec{\nabla}\cdot\vec{B} = 0, \qquad \vec{\nabla}\times\vec{E} + \frac{\partial\vec{B}}{\partial t} = 0. \tag{13.21}$$

Note that the Maxwell equation (13.18) can be solved by expressing $\vec{B}$ in terms of the magnetic vector potential $\vec{A}(\vec{r}, t)$ as follows

$$\vec{B} = \vec{\nabla}\times\vec{A} \qquad \Longrightarrow \qquad \vec{\nabla}\cdot\vec{B} = 0, \tag{13.22}$$

where we have used the vector identity which states that the divergence of a curl is always zero, $\vec{\nabla}\cdot\left(\vec{\nabla}\times\vec{A}\right) = 0$. Inserting $\vec{B} = \vec{\nabla}\times\vec{A}$ into (13.19), we obtain

$$\vec{\nabla}\times\left(\vec{E} + \frac{\partial\vec{A}}{\partial t}\right) = 0. \tag{13.23}$$

Using another identity from vector calculus where the curl of a gradient is always zero, we can express (13.23) as the gradient of a scalar, electric potential $U(\vec{r}, t)$ as follows

$$\vec{\nabla}\times\left(\vec{E} + \frac{\partial\vec{A}}{\partial t}\right) = 0 \quad \Longrightarrow \quad \vec{E} + \frac{\partial\vec{A}}{\partial t} = -\vec{\nabla}U \quad \Longrightarrow \quad \vec{E} = -\vec{\nabla}U - \frac{\partial\vec{A}}{\partial t}, \tag{13.24}$$

where we have used the fact that $\vec{\nabla} \times (\vec{\nabla} U) = 0$. So, the magnetic and electric fields that are solutions to the Maxwell equations (13.18) and (13.19) can be expressed in terms of the electric potential field, U, and the vector potential field, $\vec{A}$, as follows:

$$\vec{B} = \vec{\nabla} \times \vec{A} \qquad \vec{E} = -\nabla U - \frac{\partial \vec{A}}{\partial t}. \tag{13.25}$$

These equations, as first noted by Lorentz[4], do not determine uniquely the potentials U and $\vec{A}$. However, it is evident that the electric and magnetic fields remain the same under the following transformation

$$U \longrightarrow U' = U + \frac{\partial f}{\partial t}, \qquad \vec{A} \longrightarrow \vec{A}' = \vec{A} - \vec{\nabla} f, \tag{13.26}$$

where $f(\vec{r}, t)$ is an arbitrary function of $\vec{r}$ and t; moreover, the Maxwell equations also remain the same under this transformation. This can be seen as follows. Inserting equations (13.26) into (13.25), we can easily ascertain that $\vec{B}' = \vec{B}$ and $\vec{E}' = \vec{E}$. First, inserting $\vec{A}' = \vec{A} - \vec{\nabla} f$ into $\vec{B}' = \vec{\nabla} \times \vec{A}'$ and using the fact that the curl of the gradient of a scalar function is zero, $\vec{\nabla} \times (\vec{\nabla} f) = 0$, we see that $\vec{B}' = \vec{B}$:

$$\vec{B}' = \vec{\nabla} \times \vec{A}' = \vec{\nabla} \times (\vec{A} - \vec{\nabla} f) = \vec{\nabla} \times \vec{A} - \vec{\nabla} \times (\vec{\nabla} f) = \vec{\nabla} \times \vec{A} = \vec{B}, \tag{13.27}$$

Second, plugging $U' = U + \frac{\partial f}{\partial t}$ and $\vec{A}' = \vec{A} - \vec{\nabla} f$ into $\vec{E}' = -\nabla U' - \frac{\partial \vec{A}'}{\partial t}$, we can easily verify that $\vec{E}' = \vec{E}$:

$$\begin{aligned} \vec{E}' &= -\vec{\nabla} U' - \frac{\partial \vec{A}'}{\partial t} = -\vec{\nabla}\left(U + \frac{\partial f}{\partial t}\right) - \frac{\partial}{\partial t}\left(\vec{A} - \vec{\nabla} f\right) \\ &= \left(-\vec{\nabla} U - \frac{\partial \vec{A}}{\partial t}\right) - \vec{\nabla}\frac{\partial f}{\partial t} + \frac{\partial \vec{\nabla} f}{\partial t} \\ &= \left(-\vec{\nabla} - \frac{\partial \vec{A}}{\partial t}\right) \\ &= E, \end{aligned} \tag{13.28}$$

where we have used the fact that $\frac{\partial}{\partial t}$ and the nabla operator commute, $\frac{\partial}{\partial t}\vec{\nabla} = \vec{\nabla}\frac{\partial}{\partial t}$, since the time and space coordinates do no dependent on each other.

In view of the above, we see that the electric and magnetic potentials (U, $\vec{A}$) and their transforms (U', $\vec{A}'$) yield the same electric and magnetic fields, $\vec{E}' = \vec{E}$ and $\vec{B}' = \vec{B}$. As a result, and Since $f(\vec{r}, t)$ may be chosen to be an arbitrary function of the points $(\vec{r}, t)$, we call the transformation (13.26) a local *gauge transformation*. Since the electric and magnetic fields are invariant under this local gauge transformation, the *Maxwell formulation of electrodynamics is manifestly a gauge theory.*

[4]H.A. Lorentz, *Encyklόadie der Mathematischen Wissenschaften*, **Band V:2**, Heft 1, V. 14, 157 (1904); H.A. Lorentz, *The Theory of Electrons*, Note 5, 238, Teubner, 1909.

13.4.2 Maxwell Equations in Covariant Form

We should note that the Maxwell equations are valid in all inertial frames and are compatible with special relativity; they are covariant under Lorentz transformations. Although this Lorentz covariance is not readily obvious from the differential forms (13.16)–(13.19), we can make it more explicit by formulating the Maxwell equations in the Minkowski spacetime where space and time are treated on equal footing. This spacetime formulation will make manifest the relativistic invariance of the Maxwell equations. In this formulation, known as the covariant formulation of electrodynamics, the electric and magnetic fields are treated on equal footing; together, they form an antisymmetric second-rank tensor (see Eq .(13.35)). As we will see below, the covariant formulation of electromagnetism reduces the four Maxwell equations (13.16)–(13.19) to two (see Eq .(13.44)–(13.45)). Before dealing with the topic of covariant electrodynamics, let us recall few essential ideas about 4-tensors in spacetime.

13.4.2.1 4-Tensors in the Minkowski spacetime

In general, a 4-tensor T is a two-dimensional object that is represented in spacetime by a 4×4 matrix. Much like 4-vectors where we encounter covariant and contravariant vectors, we also have contravariant, $T^{\mu\nu}$, and covariant, $T_{\mu\nu}$, tensors, where $\mu,\ \nu = 0,\ 1,\ 2,\ 3$. By analogy to a 4-vector that has a time-like component and three space-like components, a 4-tensor has one time-like row and three space-like rows as well as one time-like column and three space-like columns. For instance, the matrix representations of the contravariant and covariant forms of the tensor F can be written as follows[5]:

$$(T^{\mu\nu}) = \begin{pmatrix} T^{00} & T^{01} & T^{02} & T^{03} \\ T^{10} & T^{11} & T^{12} & T^{13} \\ T^{20} & T^{21} & T^{22} & T^{23} \\ T^{30} & T^{31} & T^{32} & T^{33} \end{pmatrix}, \qquad \left(T_{\mu\nu}\right) = \begin{pmatrix} T_{00} & T_{01} & T_{02} & T_{03} \\ T_{10} & T_{11} & T_{12} & T_{13} \\ T_{20} & T_{21} & T_{22} & T_{23} \\ T_{30} & T_{31} & T_{32} & T_{33} \end{pmatrix} \tag{13.29}$$

Just as in the case of 4-vectors where the components of a covariant vector can be obtained from those of its contravariant vector by simply flipping the signs of the space components of the contravariant components (i.e., $x_\mu = g_{\mu\nu}x^\nu$), we can also obtain the covariant form of the tensor T in terms of the Minkowski metric tensor g as follows:

$$\begin{aligned} \left(T_{\mu\nu}\right) &= \left(g_{\mu\lambda}\right)\left(F^{\lambda\delta}\right)\left(g_{\delta\nu}\right) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} T^{00} & T^{01} & T^{02} & T^{03} \\ T^{10} & T^{11} & T^{12} & T^{13} \\ T^{20} & T^{21} & T^{22} & T^{23} \\ T^{30} & T^{31} & T^{32} & T^{33} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} T^{00} & -T^{01} & -T^{02} & -T^{03} \\ -T^{10} & T^{11} & T^{12} & T^{13} \\ -T^{20} & T^{21} & T^{22} & T^{23} \\ -T^{30} & T^{31} & T^{32} & T^{33} \end{pmatrix} \end{aligned} \tag{13.30}$$

Comparing the matrices of $\left(T_{\mu\nu}\right)$ listed in (13.29) and (13.30), we see that

$$T_{00} = T^{00}, \qquad T_{0i} = -T^{0i}, \qquad T_{i0} = -T^{i0}, \qquad T_{ij} = T^{ij}, \qquad i,\ j = 1,2,3. \tag{13.31}$$

[5]We use $[T]$ or $(T^{\mu\nu})$ to denote the matrix of the tensor, while $T^{\mu\nu}$ represents the $\mu\nu$ component of the tensor T.

Additionally, much like 4-vectors, the inner product of two 4-tensors is a scalar that is the same in all reference frames. Also, by analogy to the inner product of two 4-vectors, we need to flip the signs of the space-like rows, T^{0i}, and space-like columns, T^{i0}; so, in general, the inner product of any two 4-tensors is given by

$$\begin{aligned} T_{\mu\nu}T^{\mu\nu} &= \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} T_{\mu\nu}T^{\mu\nu} = T_{00}T^{00} + T_{0i}T^{0i} + T_{i0}T^{i0} + T_{ij}T^{ij} \\ &= T^{00}T^{00} - T^{0i}T^{0i} - T^{i0}T^{i0} + T^{ij}T^{ij}. \end{aligned} \tag{13.32}$$

Example 13.1
Show that any second-rank tensor T can be written as the sum of a symmetric tensor, T_S, and antisymmetric tensor, T_A.

Solution
First, the construction of $T_S^{\mu\nu}$ and $T_A^{\mu\nu}$ is self-evident:

$$T_S^{\mu\nu} = \frac{1}{2}\left(T^{\mu\nu} + T^{\nu\mu}\right), \qquad T_A^{\mu\nu} = \frac{1}{2}\left(T^{\mu\nu} - T^{\nu\mu}\right). \tag{13.33}$$

This shows that $T_S^{\mu\nu} = T_S^{\nu\mu}$ and $T_A^{\mu\nu} = -T_A^{\nu\mu}$. Adding these relations together, we find:

$$T_S^{\mu\nu} + T_A^{\mu\nu} = \frac{1}{2}\left(T^{\mu\nu} + T^{\nu\mu}\right) + \frac{1}{2}\left(T^{\mu\nu} - T^{\nu\mu}\right) = T^{\mu\nu}. \tag{13.34}$$

13.4.2.2 Maxwell Equations in Covariant Form

A distinct feature of the covariant formulation of electrodynamics is that the electric and magnetic fields are treated on equal footing; together, they form an antisymmetric second-rank tensor $F^{\mu\nu}$, called the *field strength tensor*. We have shown in Problem 13.3 how the electric and magnetic fields can be combined together to form $F^{\mu\nu}$:

$$\left(F^{\mu\nu}\right) = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}. \tag{13.35}$$

By analogy to (13.30), we can use (13.35) to obtain the covariant form of $F^{\mu\nu}$ as follows:

$$\left(F_{\mu\nu}\right) = \left(g_{\mu\lambda}\right)\left(F^{\lambda\delta}\right)\left(g_{\delta\nu}\right) = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}. \tag{13.36}$$

Comparing (13.35) and (13.36), we see that

$$F^{\mu\mu} = 0, \quad F^{01} = -\frac{E_x}{c}, \quad F^{02} = -\frac{E_y}{c}, \quad F^{03} = -\frac{E_z}{c}, \quad F^{12} = -B_z, \quad F^{13} = B_y, \quad F^{23} = -B_x, \tag{13.37}$$

$$F_{\mu\mu} = 0, \quad F_{01} = \frac{E_x}{c}, \quad F_{02} = \frac{E_y}{c}, \quad F_{03} = \frac{E_z}{c}, \quad F_{12} = -B_z, \quad F_{13} = B_y, \quad F_{23} = -B_x. \tag{13.38}$$

We can condense the previous two sets of relations as follows:

$$F_{\mu\mu} = 0, \qquad F^{0i} = -\frac{E_i}{c}, \qquad F^{ij} = \epsilon_{ijk}B_k, \qquad \mu = 0,1,2,3, \qquad i,\ j,\ k = 1,2,3, \tag{13.39}$$

$$F_{\mu\mu} = 0, \qquad F_{0i} = \frac{E_i}{c}, \qquad F_{ij} = \epsilon_{ijk}B_k, \qquad i,\ j,\ k = 1,2,3, \tag{13.40}$$

where ϵ_{ijk} is the 3-dimensional *Levi-Civita tensor*:

$$\epsilon_{ijk} = \begin{cases} -1 & \text{if } (i,\ j,\ k) \text{ is an } \textit{odd} \text{ permutation of } (1,2,3) \\ +1 & \text{if } (i,\ j,\ k) \text{ is an } \textit{even} \text{ permutation of } (1,2,3) \\ 0 & \text{for all other cases} \end{cases} \tag{13.41}$$

Using (13.31), (13.39) and (13.40) along with $F^{i0} = -F^{0i}$, we can easily verify that the inner product between two field tensors is a scalar:

$$\begin{aligned} F_{\mu\nu}F^{\mu\nu} &= F^{00}F^{00} - F^{0i}F^{0i} - F^{i0}F^{i0} + F^{ij}F^{ij} = -F^{0i}F^{0i} - F^{i0}F^{i0} + F^{ij}F^{ij} \\ &= -2F^{0i}F^{0i} + F_{ij}F^{ij} \\ &= -2F_{0i}F^{0i} + (\epsilon_{ijk}B_k)(\epsilon_{ijl}B_l) \\ &= -2(-E_i/c)(-E_i/c) + 2B_i^2 \\ &= 2\left(\vec{B}^2 - \frac{\vec{E}^2}{c^2}\right). \end{aligned} \tag{13.42}$$

Being a scalar quantity, the inner product $F_{\mu\nu}F^{\mu\nu}$ is a Lorentz invariant. Additionally, note that $F_{\mu\nu}F^{\mu\nu}$ has the physical dimensions of an energy density. It will be used below (see (13.52)) to construct the Lagrangian density of the electromagnetic field.

Next, let us see how to express field tensor $F^{\mu\nu}$ in terms of 4-differential operators, ∂_μ. Using the 4-vector potential $A^\mu = (A^0, \vec{A}) = \left(U/c, \vec{A}\right)$, we have shown in Problem 13.3 how the two equations (13.25) lead to the following differential and matrix forms of the field strength tensor:

$$\boxed{F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \equiv \frac{\partial A^\nu}{\partial x_\mu} - \frac{\partial A^\mu}{\partial x_\nu} \quad \Longrightarrow \quad (F^{\mu\nu}) = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}.} \tag{13.43}$$

Let us now see how to express the Maxwell equations in terms of the electromagnetic field tensor $F^{\mu\nu}$. As shown in Problem 13.4, while the inhomogeneous Maxwell equations (13.16) and (13.17) can be written in *Lorentz covariant form* by means of $F^{\mu\nu}$ as

$$\boxed{\partial_\mu F^{\mu\nu} = \mu_0 j^\nu,} \tag{13.44}$$

the homogeneous (13.18) and (13.19) can be can be expressed in covariant form as follows

$$\boxed{\partial_\lambda F^{\mu\nu} + \partial_\mu F^{\nu\lambda} + \partial_\nu F^{\lambda\mu} = 0.} \tag{13.45}$$

This is known as the *Bianchi identity* for the field strength tensor $F^{\mu\nu}$. Note that, in a source-free region of space, where no charge density nor any current density are present, $j^\mu = 0$, Equation (13.44) reduces to:

$$\partial_\mu F^{\mu\nu} = 0. \tag{13.46}$$

In summary, by rewriting and reducing the Maxwell equations (13.16)–(13.19) to the covariant forms (13.44)–(13.45), it is self-evident they are form-invariant with respect to Lorentz transformations; they have the same form in all inertial frames. Hence, they are compatible with special relativity.

Example 13.2
Using the covariant form of the Maxwell equations, $\partial_\mu F^{\mu\nu} = \mu_0 j^\nu$, show that $\partial_\mu j^\mu = 0$.

Solution
Applying ∂_ν to the Maxwell equation, $\partial_\mu F^{\mu\nu} = \mu_0 j^\nu$, we obtain

$$\partial_\nu \partial_\mu F^{\mu\nu} = \mu_0 \partial_\nu j^\nu. \tag{13.47}$$

On the one hand, since the product $\partial_\nu \partial_\mu$ is symmetric, $\partial_\nu \partial_\mu = \partial_\mu \partial_\nu$, and since $F^{\mu\nu}$ is antisymmetric, $F^{\mu\nu} = -F^{\nu\mu}$, we can write the left-hand side of (13.47) as

$$\partial_\nu \partial_\mu F^{\mu\nu} = -\partial_\mu \partial_\nu F^{\nu\mu}. \tag{13.48}$$

Now, since the summation indices μ and ν are dummy variables, we can rename them to ν and μ, respectively, in the left-hand side of (13.48) while keeping the right-hand side intact:

$$\partial_\mu \partial_\nu F^{\nu\mu} = -\partial_\mu \partial_\nu F^{\nu\mu} \qquad \Longrightarrow \qquad \partial_\mu \partial_\nu F^{\nu\mu} = 0. \tag{13.49}$$

We have actually obtained this result in Problem (13.2) where we have used a more general approach to prove that the inner product between a symmetric tensor $T_S{}^{\mu\nu}$ and an antisymmeteric tensor $T_A^{\mu\nu}$ is zero: $T_{S\mu\nu} T_A^{\mu\nu} = 0$ (see Eq. (13.141)).

Finally, inserting (13.49) into the left-hand side of (13.47), we obtain

$$\partial_\nu j^\nu = 0. \tag{13.50}$$

This equation is the continuity condition:

$$\partial_\nu j^\nu = 0 \qquad \Longrightarrow \qquad \frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0, \tag{13.51}$$

where we have used the relations $\partial_\mu = \left(\partial_0, \vec{\nabla}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla}\right)$ and the 4-vector current density $j^\mu = (c\rho, \vec{j})$.

13.4.3 Lagrangian Derivation of the Maxwell Equations

In what follows, we are going to show how to derive the *homogeneous* and then the *inhomogeneous* Maxwell equations, (13.46) and (13.44), using the *Lagrangian formulation*. The real issue now is how to construct the Lagrangian density for the electromagnetic field? In classical field theory, and as seen in (13.3), the Lagrangian density depends at a point $x^\mu = (ct, x, y, z)$ on the values of the fields $\phi(x^\mu)$ and their space-time derivatives $\partial_\mu\phi(x^\mu)$; i.e., $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu\phi)$. In the case of classical electromagnetism, the Lagrangian density depends on the fields A_μ and their space-time derivatives $\partial_\mu A^\mu$; so we have: $\mathcal{L} = \mathcal{L}\left(A_\mu, \partial_\mu A^\mu\right)$. In empty space where no charge density nor any current density are present, the Lagrangian density for the electromagnetic field can be expressed in terms of the field strength tensor $F^{\mu\nu}$ as follows:

$$\mathcal{L}\left(A^\mu, \partial_\mu A^\mu\right) = -\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu}, \tag{13.52}$$

where the dependence of $F^{\mu\nu}$ on A^μ and $\partial_\mu A^\mu$ is shown in (13.43); actually, $F^{\mu\nu}$ depends only on $\partial_\mu A^\mu$, not on A^μ. We should note that the choice of (13.52) is not random; it is justified by the facts that the Lagrangian density is a scalar quantity and that it must have the dimensions of an energy density. As shown in (13.42), the inner product $F^{\mu\nu}F_{\mu\nu}$ satisfies both of these two conditions: it is a scalar and it has the dimensions of an energy density.

Using the differential form of $F^{\mu\nu}$ listed in (13.43), we can show that

$$\begin{aligned} F_{\mu\nu}F^{\mu\nu} &= \left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) \\ &= \left(\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu\right) - \left(\partial_\mu A_\nu\right)\left(\partial^\nu A^\mu\right) - \left(\partial_\nu A_\mu\right)\left(\partial^\mu A^\nu\right) + \left(\partial_\nu A_\mu\right)\left(\partial^\nu A^\mu\right) \\ &= 2\left(\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu\right) - 2\left(\partial_\mu A_\nu\right)\left(\partial^\nu A^\mu\right), \end{aligned} \tag{13.53}$$

where, in the last step, we have used the facts that $\left(\partial_\nu A_\mu\right)\left(\partial^\nu A^\mu\right) = \left(\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu\right)$ and $\left(\partial_\nu A_\mu\right)\left(\partial^\mu A^\nu\right) = \left(\partial_\mu A_\nu\right)\left(\partial^\nu A^\mu\right)$, since μ and ν are dummy variables and, hence, we can interchange them.

Inserting (13.53) into (13.52), we can write the Lagrangian density of an electromagnetic field A_μ in empty space as follows:

$$\mathcal{L}\left(A_\mu, \partial_\mu A^\mu\right) = -\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{2\mu_0}\left[\left(\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu\right) - \left(\partial_\mu A_\nu\right)\left(\partial^\nu A^\mu\right)\right]. \tag{13.54}$$

Now that we have specified the Lagrangian density, let us apply the Euler–Lagrange equation to (13.54):

$$\frac{\partial\mathcal{L}}{\partial A_\nu} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu A_\nu\right)}\right) = 0. \tag{13.55}$$

This reduces to finding the derivatives $\frac{\partial\mathcal{L}}{\partial A_\nu}$ and $\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu A_\nu\right)}$. First, since $\mathcal{L}$ as given by (13.54) does not depend on A_ν (it depends on $\partial_\mu A_\nu$ only), we obtain at once:

$$\frac{\partial\mathcal{L}}{\partial A_\nu} = 0. \tag{13.56}$$

Next, using (13.54), we can easily obtain the other derivative

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu A_\nu\right)} &= -\frac{1}{2\mu_0}\frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left[\left(\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu\right)-\left(\partial_\mu A_\nu\right)\left(\partial^\nu A^\mu\right)\right] \\ &= -\frac{1}{\mu_0}\left(\partial^\mu A^\nu-\partial^\nu A^\mu\right) \\ &= -\frac{1}{\mu_0}F^{\mu\nu}. \end{aligned} \tag{13.57}$$

Finally, plugging (13.56) and (13.57) into (13.55), we obtain

$$\partial_\nu F^{\mu\nu}=0. \tag{13.58}$$

This is the sourceless or homogeneous Maxwell equations (13.46) that we have derived from covariant electrodynamics. So, by using the Lagrangian density (13.52), the Euler–Lagrange equation gave us back the Maxwell's equations; hence, the choice $\mathcal{L}=-\frac{1}{4\mu_0}F^{\mu\nu}F_{\mu\nu}$ is justified.

Lagrangian Derivation of the Inhomogeneous Maxwell Equations

Now, let us look at the Lagrangian derivation of the *inhomogeneous* Maxwell equations (13.44). In case where there is a source of charge ρ and a 3-current $\vec{J}$ (i.e., a 4-current density $j_\nu=\left(c\rho,\ \vec{J}\right)$), we would have a term that accounts for the interaction between the j^ν source and the electromagnetic field A_ν:

$$\mathcal{L}_{int}=-j^\nu A_\nu. \tag{13.59}$$

This interaction term is a scalar and has the dimensions of an energy density. In this case, we can construct a Lagrangian density by adding (13.59) to the Lagrangian density for the electromagnetic field (13.52):

$$\mathcal{L}\left(A_\mu,\ \partial_\mu A^\mu\right)=-\frac{1}{4\mu_0}F^{\mu\nu}F_{\mu\nu}-j^\nu A_\nu. \tag{13.60}$$

In working with this Lagrangian, we will treat each component of A_μ as an independent field.

To apply the Euler–Lagrange equation (13.55) to the Lagrangian density (13.60), we need to obtain first the derivative $\frac{\partial \mathcal{L}}{\partial A_\nu}$. Since the dependence of (13.60) on A_ν is contained in the term $-j^\nu A_\nu$, we can easily write

$$\frac{\partial \mathcal{L}}{\partial A_\nu}=\frac{\partial\left(-j^\nu A_\nu\right)}{\partial A_\nu}=-j^\nu. \tag{13.61}$$

Finally, inserting (13.57) and (13.61) into the Euler–Lagrange equation (13.55), we recover the covariant form of the inhomogeneous Maxwell equations:

$$\partial_\mu F^{\mu\nu}=\mu_0 j^\nu. \tag{13.62}$$

This is the same equation we have obtained above (see (13.44)) from the covariant formulation of electromagnetism.

13.5 Lagrangian Derivation of the Klein-Gordon Equation

We have seen in Chapter 12 that, when the solution ϕ of the Klein-Gordon equation is viewed as a single-particle wave function, we run into several untractable problems, chiefly among them: the probability density is *negative*. We also saw that this problem was not resolved until the Klein-Gordon equation was re-interpreted as a *field equation* for a *spinless*, or *scalar*[6], particle. So, in what follows, we are going to consider $\phi(x^\mu)$ to be a real *scalar field*, not a single-particle wave function.

The most important task of this section is to construct a Lagrangian density for the field $\phi(x^\mu)$ from which we can derive the Klein-Gordon equation using the least action principle; this equation will be a *field equation*, not a single-particle equation.

Since a particle is represented in field theory by excitations of the field, we assume that its corresponding field $\phi(x^\mu)$ is defined everywhere in space-time. In general, fields change under rotations and Lorentz boosts (i.e., Lorentz transformations). However, scalar fields have the distinctive property of remaining invariant under Lorentz transformations: $\phi(x^\mu) = \phi(x'^\mu)$. Spinless particles are called scalar particles, since their fields are Lorentz invariant; they remain the same in all inertial frames. In field theory, the Lagrangian governing the dynamics of the system needs to be Lorentz invariant. As a result, when we get to the task of constructing the Klein-Gordon and Dirac Lagrangians below, we will make sure that they are invariant under Lorentz transformations.

As mentioned above, the Lagrangian formalism is an effective theoretical framework for dealing with field theories. As such, we are going to discuss how to obtain the Klein-Gordon Lagrangian density for a scalar field, then apply it to the Euler–Lagrange equation and then show how to recover the Klein-Gordon equation.

The Klein-Gordon Lagrangian density for a free real scalar field ϕ (a single component function on space-time) is given by

$$\mathcal{L}(\phi,\ \partial_\mu\phi) = \frac{1}{2}\left(\partial_\mu\phi\partial^\mu\phi - \frac{m^2c^2}{\hbar^2}\phi^2\right). \tag{13.63}$$

Note that $\mathcal{L}$ depends on the field ϕ and its derivative $\partial_\mu\phi$, but not explicitly on the coordinates x^μ; if $\mathcal{L}$ depends explicitly on x^μ, it will violate the Lorentz invariance. That is, if $\mathcal{L}$ is to be invariant under space-time translations, then it cannot depend explicitly on x^μ.

An important question arises at this level: Where did this Lagrangian density come from? Can it be derived from first principles? The simple answer is no! It was postulated. However, although we have postulated it, we were guided by two major requirements:

- We needed to incorporate all the symmetries we know about the system into the Lagrangian density. For instance, since the Klein-Gordon field is a scalar, we would need to construct a Lagrangian density that is a Lorentz scalar; we want the Lagrangian density to keep the same form in all inertial frames. The simplest scalars that can be constructed out of the field, ϕ, and its first space-time derivative, $\partial_\mu\phi$, are ϕ^2 and $\partial_\mu\phi\partial^\mu\phi$, respectively. The terms $\partial_\mu\phi\partial^\mu\phi$ represents the kinetic energy density and $\frac{m^2c^2}{\hbar^2}\phi^2$ is the rest mass-energy density; we will deal with the physical dimensions of these terms below (see Eq. (13.74)). Additionally, since we want to end up with a linear equation (the

[6] A scalar particle is a particle that has no spin and no internal structure.

Euler–Lagrange equation), higher powers of the field (i.e., higher than ϕ^2 and $\partial_\mu\phi\partial^\mu\phi$) would need to be avoided in the Lagrangian density.

- At the end, we wanted to be able to recover the Klein-Gordon equation.

Guided by these two essential requirements, we were able to obtain the form (13.63) up to some inconsequential[7] constants; so this is guided intuition. More precisely, the form (13.63) was intuited by analogy to the classical Lagrangian, $L = T - V$, after making use of the substitutions (13.1). So, the Klein-Gordon Lagrangian density (13.63) has the same structure as the classical Lagrangian except that we have no potential energy since we are dealing with a *free* field.

One should recall that the Lagrangian density cannot be derived from first principles, but once it is postulated, the rest will be on autopilot; i.e., everything will follow its logical mathematical course until we obtain the Klein-Gordon equation. Klein and Gordon have not obtained their equation from first principles either; they have started with the Schrödinger equation, which in itself was also postulated. Bottom line, we have to start somewhere.

Derivation of the Klein-Gordon equation from the Euler–Lagrange equation
Now, to obtain the Klein-Gordon equation, we need simply to apply the Lagrangian density (13.63) to the Euler–Lagrange equation of motion (13.15). For this, we would need to find the derivatives $\frac{\partial\mathcal{L}}{\partial\phi}$ and $\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}$. First, we can easily write the first derivative:

$$\frac{\partial\mathcal{L}}{\partial\phi} = \frac{1}{2}\frac{\partial}{\partial\phi}\left(\partial_\mu\phi\partial^\mu\phi - \frac{m^2c^2}{\hbar^2}\phi^2\right) = -\frac{m^2c^2}{\hbar^2}\phi. \tag{13.64}$$

Next, using the fact that

$$\partial_\mu\phi\partial^\mu\phi = \partial^0\phi\partial^0\phi - \partial^i\phi\partial^i\phi \equiv \left(\partial^0\phi\right)^2 - \left(\partial^i\phi\right)^2, \tag{13.65}$$

we can write the Lagrangian density (13.63) as

$$\mathcal{L}(\phi,\, \partial_\mu\phi) = \frac{1}{2}\left[\left(\partial^0\phi\right)^2 - \left(\partial^i\phi\right)^2 - \frac{m^2c^2}{\hbar^2}\phi^2\right]. \tag{13.66}$$

Since $\partial_0 \equiv \partial^0$, we have

$$\frac{\partial\mathcal{L}}{\partial\left(\partial_0\phi\right)} \equiv \frac{\partial\mathcal{L}}{\partial\left(\partial^0\phi\right)} = \frac{1}{2}\frac{\partial}{\partial\left(\partial^0\phi\right)}\left[\left(\partial^0\phi\right)^2 - \left(\partial^i\phi\right)^2 - \frac{m^2c^2}{\hbar^2}\phi^2\right] = \partial^0\phi, \tag{13.67}$$

and since $\partial_j \equiv -\partial^j$, where $j = 1,\, 2,\, 3$, we can obtain the following derivative

$$\frac{\partial\mathcal{L}}{\partial\left(\partial_j\phi\right)} \equiv -\frac{\partial\mathcal{L}}{\partial\left(\partial^i\phi\right)} = -\frac{1}{2}\frac{\partial}{\partial\left(\partial^i\phi\right)}\left[\left(\partial^0\phi\right)^2 - \left(\partial^i\phi\right)^2 - \frac{m^2c^2}{\hbar^2}\phi^2\right] = \partial^i\phi. \tag{13.68}$$

We can easily condense (13.67) and (13.68) into

$$\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)} = \partial^\mu\phi \qquad \Longrightarrow \qquad \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\right) = \partial_\mu\partial^\mu\phi. \tag{13.69}$$

[7]The constants in (13.63) were chosen in such a way as to recover the Klein-Gordon equation at the end.

Finally, substituting (13.64) and (13.69) into (13.15), we obtain

$$\boxed{\left(\partial_\mu \partial^\mu + \frac{m^2c^2}{\hbar^2}\right)\phi = 0.} \tag{13.70}$$

This equation is the Klein-Gordon equation we have derived in Chapter 12. At this level, it is instructive to look back and contrast the two approaches used to derive the Klein-Gordon equation: while Chapter 12 follows the traditional, single-particle approach which starts from the Schrödinger equation and then uses the energy expression $E^2 = \vec{p}^2c^2 + m^2c^4$, Subsection 13.5 offers a more formal method that rests on a variational principle that, after postulating the Lagrangian density, leads ultimately to the Euler–Lagrange equation.

Note that, and as mentioned above, the Lagrangian density (13.63) is not the only one that can lead to the Klein-Gordon equation. For instance, using the following Lagrangian density for a *complex*, spin-zero field

$$\mathcal{L}(\phi,\, \partial_\mu\phi,\, \phi^*,\, \partial_\mu\phi^*) = \frac{1}{2}\left[\left(\partial_\mu\phi\right)^*\left(\partial^\mu\phi\right) - \frac{m^2c^2}{\hbar^2}\phi^*\phi\right], \tag{13.71}$$

we can show that it will lead to the Klein-Gordon equation (13.70); see Example 13.3 for an illustration.

Physical dimensions of the quantities encountered above

Let us take a short detour to discuss the physical dimensions (in SI units) of the quantities involved in the formalism outlined above such as the Lagrangian, Lagrangian density, the field ϕ, and the action. First, since the Lagrangian has the dimensions of energy, we have

$$[\text{Lagrangian}] = \text{ML}^2\text{T}^{-2}, \tag{13.72}$$

where M = *Mass*, L = *Length*, and T = *Time*. Next, since the Lagrangian is the volume integration of the Lagrangian density $\mathcal{L}$ (i.e., $L = \int d^3x\mathcal{L}$), the Lagrangian density is bound to have the dimensions of *energy per unit volume*:

$$[\text{Lagrangian}] = [\mathcal{L}] \times \text{L}^3 \implies \text{ML}^2\text{T}^{-2} = [\mathcal{L}] \times \text{L}^3 \implies \boxed{[\mathcal{L}] = ML^{-1}T^{-2}}. \tag{13.73}$$

To find the dimensions of the Klein-Gordon field ϕ, we need to make use of equations (13.63) and (13.73); that is, from (13.63), we see that $[\mathcal{L}] = [\partial_\mu\phi] \times [\partial^\mu\phi] \equiv [\partial^2][\phi^2]$ and since the physical dimension of the operator ∂ is the inverse of a length, $[\partial] = L^{-1}$, we can write

$$[\mathcal{L}] = [\partial^2][\phi^2] \implies ML^{-1}T^{-2} = L^{-2}[\phi^2] \implies [\phi^2] = MLT^{-2} \implies \boxed{[\phi] = M^{1/2}L^{1/2}T^{-1}}. \tag{13.74}$$

Now, since the action is given by $S = \int dtL$, the dimensions of the action are those of a Lagrangian multiplied by time; hence, using (13.72), we have:

$$[S] = [\text{Lagrangian}] \times \text{T} \equiv [\text{energy}] \times \text{T} \implies [\text{S}] = \text{ML}^2\text{T}^{-2} \times \text{T} \implies \boxed{[S] = ML^2T^{-1}}. \tag{13.75}$$

So, as expected, the action S has the same dimensions as the reduced Planck constant, $\hbar$.

Example 13.3
Consider the Lagrangian density for a complex, spin-zero field that is listed in Eq. (13.71).
(a) Show that the Euler–Lagrange equation for ϕ is the Klein-Gordon equation for ϕ^*.
(b) Show that the Euler–Lagrange equation for ϕ^* is the Klein-Gordon equation for ϕ.

Solution
(a) To find the Euler–Lagrange equations for ϕ with respect to the Lagrangian density (13.71), we need to begin by finding the derivatives $\frac{\partial \mathcal{L}}{\partial \phi}$ and $\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}$:

$$\frac{\partial \mathcal{L}}{\partial \phi} = \frac{1}{2}\frac{\partial}{\partial \phi}\left[\left(\partial_\mu \phi\right)^*\left(\partial^\mu \phi\right) - \frac{m^2c^2}{\hbar^2}\phi^*\phi\right] = -\frac{m^2c^2}{\hbar^2}\phi^*. \tag{13.76}$$

Using the fact that

$$\left(\partial_\mu \phi\right)^*\left(\partial^\mu \phi\right) \equiv \partial_\mu \phi^* \partial^\mu \phi = \partial^0 \phi^* \partial^0 \phi - \partial^j \phi^* \partial^j \phi, \tag{13.77}$$

we can write the Lagrangian density $\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi\right)^*\left(\partial^\mu \phi\right) - \frac{m^2c^2}{2\hbar^2}\phi^*\phi$ as

$$\mathcal{L} = \frac{1}{2}\left[\partial^0 \phi^* \partial^0 \phi - \partial^j \phi^* \partial^j \phi - \left(\frac{mc}{\hbar}\right)^2 \phi^*\phi\right]. \tag{13.78}$$

Next, since $\partial_0 \equiv \partial^0$, we can obtain the following derivative

$$\frac{\partial \mathcal{L}}{\partial\left(\partial_0 \phi\right)} \equiv \frac{\partial \mathcal{L}}{\partial\left(\partial^0 \phi\right)} = \frac{1}{2}\frac{\partial}{\partial\left(\partial^0 \phi\right)}\left[\partial^0 \phi^* \partial^0 \phi - \partial^j \phi^* \partial^j \phi - \left(\frac{mc}{\hbar}\right)^2 \phi^*\phi\right] = \partial^0 \phi^*, \tag{13.79}$$

and since $\partial_j \equiv -\partial^j$, where $j = 1, 2, 3$, we can also obtain the following derivative

$$\frac{\partial \mathcal{L}}{\partial\left(\partial_j \phi\right)} \equiv -\frac{\partial \mathcal{L}}{\partial\left(\partial^j \phi\right)} = -\frac{1}{2}\frac{\partial}{\partial\left(\partial^j \phi\right)}\left[\partial^0 \phi^* \partial^0 \phi - \partial^j \phi^* \partial^j \phi - \left(\frac{mc}{\hbar}\right)^2 \phi^*\phi\right] = \partial^j \phi^*. \tag{13.80}$$

We can easily condense (13.79) and (13.80) into

$$\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} = \partial^\mu \phi^* \qquad \Longrightarrow \qquad \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)}\right) = \partial_\mu \partial^\mu \phi^*. \tag{13.81}$$

Finally, substituting (13.76) and (13.81) into the Euler–Lagrange equation (13.15), we obtain

$$\left(\partial_\mu \partial^\mu + \frac{m^2c^2}{\hbar^2}\right)\phi^* = 0. \tag{13.82}$$

This equation is the Klein-Gordon equation for ϕ^*; see (13.70).

(b) Second, let us find the Euler–Lagrange equations for ϕ^*. For this, we need to begin by finding the derivatives $\frac{\partial \mathcal{L}}{\partial \phi^*}$ and $\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^*)}$:

$$\frac{\partial \mathcal{L}}{\partial \phi^*} = \frac{1}{2}\frac{\partial}{\partial \phi^*}\left[\left(\partial_\mu \phi\right)^*\left(\partial^\mu \phi\right) - \frac{m^2c^2}{\hbar^2}\phi^*\phi\right] = -\frac{m^2c^2}{\hbar^2}\phi. \tag{13.83}$$

Using (13.78) and the fact that $\partial_0 \equiv \partial^0$, we have

$$\frac{\partial \mathcal{L}}{\partial\left(\partial_0 \phi^*\right)} \equiv \frac{\partial \mathcal{L}}{\partial\left(\partial^0 \phi^*\right)} = \frac{1}{2} \frac{\partial}{\partial\left(\partial^0 \phi^*\right)}\left[\partial^0 \phi^* \partial^0 \phi - \partial^j \phi^* \partial^j \phi - \left(\frac{mc}{\hbar}\right)^2 \phi^* \phi\right] = \partial^0 \phi. \tag{13.84}$$

Next, since $\partial_j \equiv -\partial^j$, and taking the derivative of (13.78) with respect to $(\partial_j \phi^*)$, we obtain

$$\frac{\partial \mathcal{L}}{\partial\left(\partial_j \phi^*\right)} \equiv -\frac{\partial \mathcal{L}}{\partial\left(\partial^j \phi^*\right)} = -\frac{1}{2} \frac{\partial}{\partial\left(\partial^j \phi^*\right)}\left[\partial^0 \phi^* \partial^0 \phi - \partial^j \phi^* \partial^j \phi - \left(\frac{mc}{\hbar}\right)^2 \phi^* \phi\right] = \partial^j \phi. \tag{13.85}$$

We now can condense (13.84) and (13.85) into

$$\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi^*\right)} = \partial^\mu \phi \qquad \Longrightarrow \qquad \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi^*\right)}\right) = \partial_\mu \partial^\mu \phi. \tag{13.86}$$

Finally, substituting (13.83) and (13.86) into the Euler–Lagrange equation (13.15), we obtain

$$\left(\partial_\mu \partial^\mu + \frac{m^2 c^2}{\hbar^2}\right) \phi = 0. \tag{13.87}$$

This equation is the Klein-Gordon equation for ϕ.

In summary, we conclude that by applying the Euler–Lagrange equation on the Lagrangian density $\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi\right)^*\left(\partial^\mu \phi\right) - \frac{m^2 c^2}{2\hbar^2} \phi^* \phi$, we obtain two separate Klein-Gordon equations, one for ϕ and the other for ϕ^*; these two equations are complex conjugates of each other; namely, (13.82) is nothing but the complex conjugate of (13.87).

13.6 Lagrangian Derivation of the Dirac Equation

Following the same method used above to obtain the Klein-Gordon Lagrangian density, we are going to construct a Lorentz-invariant Lagrangian density that leads to the Dirac equation. To ensure the Lorentz invariance of the Dirac Lagrangian density, we are going to construct it using quantities that are scalar such as $\overline{\psi}\psi$ and $\overline{\psi}\gamma^\mu \partial_\mu \psi$, where $\overline{\psi}$ is the adjoint spinor field (i.e., $\overline{\psi} \equiv \psi^\dagger \gamma^0$) of the Dirac spinor, spin-1/2 field ψ.

The *Dirac Lagrangian density* for a spinor field ψ (a four-component complex spinor representing a spin 1/2 field) is postulated to be given by

$$\mathcal{L}(\psi, \partial_\mu \psi, \overline{\psi}, \partial_\mu \overline{\psi}) = i\hbar c \overline{\psi} \gamma^\mu \partial_\mu \psi - mc^2 \overline{\psi} \psi, \tag{13.88}$$

where ψ and $\overline{\psi}$ are the independent field variables and where $i\hbar c \overline{\psi} \gamma^\mu \partial_\mu \psi$ represents the kinetic energy density; the spin (i.e., spin 1/2) is embedded within the γ^μ matrix. As for the term $mc^2 \overline{\psi} \psi$, it represents the rest mass-energy density. The Lagrangian density (13.88) describes a spin 1/2 *free* field, ψ, with no interaction term with any other field or particle.

We should note that, as shown in Equation (13.73), since the Lagrangian density has the dimensions of energy per unit volume and, hence, are given in SI units by $[\mathcal{L}] = ML^{-1}T^{-2}$, the dimensions of the Dirac field, ψ, can be inferred from (13.88) as follows

$$[\mathcal{L}] \equiv [m] \times [c^2] \times [\psi^2] \implies ML^{-1}T^{-2} = ML^2T^{-2}[\psi^2] \implies \boxed{[\psi] = L^{-3/2}}. \tag{13.89}$$

That is, the Dirac field has dimensions of $[\psi] = [\text{Length}]^{-3/2}$ or of the square root of a volume.

Now, let us look at the Euler–Lagrange equation (13.15) for the Dirac spinor ψ

$$\frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} \right) = 0. \tag{13.90}$$

Using (13.88), we can obtain the first and second terms of (13.90) as follows:

$$\frac{\partial \mathcal{L}}{\partial \psi} = \frac{\partial}{\partial \psi} \left(i\hbar c \overline{\psi} \gamma^\mu \partial_\mu \psi - mc^2 \overline{\psi} \psi \right) \equiv \frac{\partial}{\partial \psi} \left(-mc^2 \overline{\psi} \psi \right) = -mc^2 \overline{\psi}, \tag{13.91}$$

and

$$\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} = \frac{\partial \left(i\hbar c \overline{\psi} \gamma^\mu \partial_\mu \psi - mc^2 \overline{\psi} \psi \right)}{\partial (\partial_\mu \psi)} \equiv \frac{\partial \left(i\hbar c \overline{\psi} \gamma^\mu \partial_\mu \psi \right)}{\partial (\partial_\mu \psi)} = i\hbar c \overline{\psi} \gamma^\mu. \tag{13.92}$$

This last equation leads to

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} \right) = i\hbar c \left(\partial_\mu \overline{\psi} \right) \gamma^\mu. \tag{13.93}$$

Next, substituting (13.91) and (13.93) into (13.90), we obtain

$$\boxed{i\hbar \left(\partial_\mu \overline{\psi} \right) \gamma^\mu + mc \overline{\psi} = 0.} \tag{13.94}$$

This equation is the adjoint of the Dirac equation derived in Chapter 12; it is the Dirac equation for *antiparticles*.

Second, combining the Euler–Lagrange equation for the adjoint spinor $\overline{\psi}$

$$\frac{\partial \mathcal{L}}{\partial \overline{\psi}} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \overline{\psi})} \right) = 0, \tag{13.95}$$

with the Lagrangian density (13.88), we see that the first term of (13.95) is given by

$$\frac{\partial \mathcal{L}}{\partial \overline{\psi}} = \frac{\partial}{\partial \overline{\psi}} \left(i\hbar c \overline{\psi} \gamma^\mu \partial_\mu \psi - mc^2 \overline{\psi} \psi \right) = i\hbar c \gamma^\mu \partial_\mu \psi - mc^2 \psi, \tag{13.96}$$

and the second term is zero because $\mathcal{L}$ does not depend on $\partial_\mu \overline{\psi}$:

$$\frac{\partial \mathcal{L}}{\partial (\partial_\mu \overline{\psi})} = \frac{\partial}{\partial (\partial_\mu \overline{\psi})} \left(i\hbar c \overline{\psi} \gamma^\mu \partial_\mu \psi - mc^2 \overline{\psi} \psi \right) = 0; \tag{13.97}$$

Next, substituting (13.96) and (13.97) into (13.95), we obtain

$$i\hbar c\gamma^\mu \partial_\mu \psi - mc^2\psi = 0, \tag{13.98}$$

which can be rewritten as

$$\boxed{\left(i\hbar\gamma^\mu \partial_\mu - mc\right)\psi = 0.} \tag{13.99}$$

This equation is the Dirac equation for *particles* that we have derived in Chapter 12.

Remarks

- Very much like classical Lagrangians, the field Lagrangian densities of Klein-Gordon and Dirac, (13.63) and (13.88), are not unique. For instance, when we multiply them by constants and/or add constants to them, we would end up with the same field equations because the terms generated by these constants would cancel out in the Euler–Lagrange equation; that is, $\mathcal{L}$ and $\mathcal{L}' = a\mathcal{L} + b$, where a and b are constants, would lead to the same field equations. Similarly, we can also show the invariance of the field equations under a shift of the Lagrangian densities by a total time derivative of a function: $\mathcal{L}$ and $\mathcal{L}' = \mathcal{L} + df(t)/dt$, where $f(t)$ is a time-dependent function, would produce the same field equations because $f(t)$ vanishes at the end points, $f(t_1) = f(t_2) = 0$.

- We can construct the Dirac Lagrangian density for a spin 1/2 particle of charge q that moves in the presence of an electromagnetic A^μ by simply using the minimal substitution relations $i\hbar\partial^\mu \longrightarrow i\hbar\partial^\mu - qA^\mu$ in the Dirac Lagrangian density (13.88):

$$\mathcal{L} = i\hbar c\overline{\psi}\gamma_\mu\partial^\mu\psi - mc^2\overline{\psi}\psi \quad \longrightarrow \quad \mathcal{L} = c\overline{\psi}\gamma_\mu\left(i\hbar\partial^\mu - qA^\mu\right)\psi - mc^2\overline{\psi}\psi. \tag{13.100}$$

  As shown in Problem 13.6 on Page 931, the Dirac equation for a spin 1/2 charged particle q in an electromagnetic field A^μ is given by

$$i\hbar\frac{\partial\psi}{\partial t} = \left[c\vec{\alpha}\cdot\left(\hat{\vec{p}} - q\vec{A}\right) + mc^2\beta + \hat{V}\right]\psi, \tag{13.101}$$

  This is the same equation we have derived in Chapter 12.

- As mentioned above, the Lagrangian density (13.88) is not unique; for instance, and as shown in Example (13.4) below, we can obtain the Dirac equation by using the following form that includes ψ and $\overline{\psi}$ symmetrically

$$\mathcal{L}(\psi, \partial_\mu\psi, \overline{\psi}, \partial_\mu\overline{\psi}) = \frac{i}{2}\hbar c\left[\overline{\psi}\gamma^\mu\left(\partial_\mu\psi\right) - \left(\partial_\mu\overline{\psi}\right)\gamma^\mu\psi\right] - mc^2\overline{\psi}\psi. \tag{13.102}$$

Example 13.4

Consider the Lagrangian density for a spin-1/2 field that is listed in Eq. (13.102).

(a) Show that the Euler–Lagrange equation for ψ is the Dirac equation for $\overline{\psi}$.

(b) Show that the Euler–Lagrange equation for $\overline{\psi}$ is the Dirac equation for ψ.

Solution

(a) To find the Euler–Lagrange equations for ψ with respect to the Lagrangian density (13.102), we need to begin by finding the following two derivatives:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \psi} &= \frac{i}{2}\hbar c \frac{\partial}{\partial \psi}\left[\overline{\psi}\gamma^{\mu}\left(\partial_{\mu}\psi\right) - \left(\partial_{\mu}\overline{\psi}\right)\gamma^{\mu}\psi\right] - mc^{2}\frac{\partial}{\partial \psi}\left(\overline{\psi}\psi\right) \\ &= \frac{i}{2}\hbar c \frac{\partial}{\partial \psi}\left[-\left(\partial_{\mu}\overline{\psi}\right)\gamma^{\mu}\psi\right] - mc^{2}\overline{\psi} \\ &= -\frac{i}{2}\hbar c\left(\partial_{\mu}\overline{\psi}\right)\gamma^{\mu} - mc^{2}\overline{\psi}, \end{aligned} \tag{13.103}$$

and

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu}\psi\right)} &= \frac{i}{2}\hbar c \frac{\partial}{\partial\left(\partial_{\mu}\psi\right)}\left[\overline{\psi}\gamma^{\mu}\left(\partial_{\mu}\psi\right) - \left(\partial_{\mu}\overline{\psi}\right)\gamma^{\mu}\psi\right] - mc^{2}\frac{\partial}{\partial\left(\partial_{\mu}\psi\right)}\left(\overline{\psi}\psi\right) \\ &= \frac{i}{2}\hbar c \frac{\partial}{\partial\left(\partial_{\mu}\psi\right)}\left[\overline{\psi}\gamma^{\mu}\left(\partial_{\mu}\psi\right)\right] \\ &= \frac{i}{2}\hbar c\,\overline{\psi}\gamma^{\mu}. \end{aligned} \tag{13.104}$$

This last equation leads to

$$\partial_{\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu}\psi\right)}\right) = \frac{i}{2}\hbar c\,\partial_{\mu}\left(\overline{\psi}\gamma^{\mu}\right). \tag{13.105}$$

Finally, inserting (13.103) and (13.105) into the Euler–Lagrange equation (13.15), we obtain

$$-\frac{i}{2}\hbar c\,\left(\partial_{\mu}\overline{\psi}\right)\gamma^{\mu} - mc^{2}\overline{\psi} - \frac{i}{2}\hbar c\overline{\psi}\gamma^{\mu} = 0, \tag{13.106}$$

which can be rewritten as

$$i\hbar\left(\partial_{\mu}\overline{\psi}\right)\gamma^{\mu} + mc\overline{\psi} = 0. \tag{13.107}$$

This is the same equation as (13.94); it is the Hermitian conjugate of the Dirac equation.

(b) Second, let us find the Euler–Lagrange equations for $\overline{\psi}$. For this, we need to find the following two derivatives of the Lagrangian density:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \overline{\psi}} &= \frac{i}{2}\hbar c \frac{\partial}{\partial \overline{\psi}}\left[\overline{\psi}\gamma^{\mu}\left(\partial_{\mu}\psi\right) - \left(\partial_{\mu}\overline{\psi}\right)\gamma^{\mu}\psi\right] - mc^{2}\frac{\partial}{\partial \overline{\psi}}\left(\overline{\psi}\psi\right) \\ &= \frac{i}{2}\hbar c \frac{\partial}{\partial \overline{\psi}}\left[\overline{\psi}\gamma^{\mu}\left(\partial_{\mu}\psi\right)\right] - mc^{2}\psi \\ &= \frac{i}{2}\hbar c\gamma^{\mu}\left(\partial_{\mu}\psi\right) - mc^{2}\psi, \end{aligned} \tag{13.108}$$

and

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \overline{\psi}\right)} &= \frac{\partial}{\partial\left(\partial_\mu \overline{\psi}\right)}\left[\frac{i}{2}\hbar c\left(\overline{\psi}\gamma^\mu\left(\partial_\mu\psi\right)-\left(\partial_\mu\overline{\psi}\right)\gamma^\mu\psi\right)-mc^2\overline{\psi}\psi\right] \\ &= \frac{\partial}{\partial\left(\partial_\mu \overline{\psi}\right)}\left[-\frac{i}{2}\hbar c\left(\partial_\mu\overline{\psi}\right)\gamma^\mu\psi\right] \\ &= -\frac{i}{2}\hbar c\gamma^\mu\psi, \end{aligned} \tag{13.109}$$

which leads to

$$\partial_\mu\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \overline{\psi}\right)}\right) = -\frac{i}{2}\hbar c\gamma^\mu\left(\partial_\mu\psi\right). \tag{13.110}$$

Now, substituting (13.108) and (13.109) into

$$\frac{\partial \mathcal{L}}{\partial \overline{\psi}} - \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \overline{\psi}\right)}\right) = 0, \tag{13.111}$$

we obtain

$$i\hbar c\gamma^\mu\partial_\mu\psi - mc^2\psi = 0, \tag{13.112}$$

which can be rewritten as

$$\left(i\hbar\gamma^\mu\partial_\mu - mc\right)\psi = 0. \tag{13.113}$$

This is the Dirac equation for *particles* that we have derived above (13.99).

In summary, we conclude that by applying the Euler–Lagrange equation on the Lagrangian density $\mathcal{L} = \frac{i}{2}\hbar c\left[\overline{\psi}\gamma^\mu\left(\partial_\mu\psi\right)-\left(\partial_\mu\overline{\psi}\right)\gamma^\mu\psi\right] - mc^2\overline{\psi}\psi$, we obtain two separate Dirac equations, one for ψ and the other for $\overline{\psi}$; these two equations are Hermitian adjoints of each other.

13.7 Beyond Classical Field Theory: Quantum Field Theory

In this chapter, we have seen how to use Lagrangians within the context of classical field theory to derive the Klein-Gordon and Dirac equations. This theory has its own limitations: it is classical. To deal with quantum phenomena, *one has to go beyond classical field theory by quantizing the classical fields*. This leads to *Quantum Field Theory*. Since a full treatment of the formalism of quantum field theory would require several chapters, we are going to mention only few essential ideas that are mainly of historical significance.

Quantum Field Theory has resulted from the combination of classical field theory, special relativity, and quantum mechanics. Its development has started in the late 1920s and continued for the rest of the twentieth century. During this period, QFT has undergone a number of formulations. Initially, it started with the description of the interaction of light (i.e., photons) with charged particles, particularly electrons; this first version of QFT is known as *quantum electrodynamics* (QED). Soon after its introduction, QED encountered a formidable challenge due to the appearance of infinities in its perturbative calculations. No sooner had

the infinities problem been resolved in the 1950s (by means of the renormalization prescription) than the second major hurdle appeared: the inability of QED to describe the strong and weak interactions. This second problem was not resolved until the introduction of *gauge field theories*, which took several decades[8] to develop. It is now understood that the fundamental interactions of nature can be described in terms of gauge theories; actually, and as seen above, since the Maxwell equations are invariant under the gauge transformation (13.26), Maxwell's electrodynamics is considered to be the first historically significant example of a gauge theory. The extensive work on gauge field theories has culminated in the development of the *Standard Model* of particle physics. This model deals with two major elements: (a) the description of the three fundamental forces of nature (i.e., the electromagnetic, weak, and strong interactions) and (b) the classification of the various elementary particles. As mentioned above, with the introduction of gauge field theories, QFT was used successfully to describe the weak and strong interactions leading to the standard model which is supported by ample experimental evidence. As a result, QFT is sometimes used synonymously with the standard model; as a reminder, the standard model, while describing accurately the electromagnetic, weak, and strong interactions, it does not cover the fourth fundamental force of gravitation which, so far, has resited quantization.

From its very inception, QFT was designed to offer a unified treatment of *particles and fields* by considering material particles to be excited states of quantum fields. Before the advent of QFT, particles and quantum fields (e.g., photons) were considered to be very different or mutually-exclusive entities. In this old view, material particles were described by means of physical states with specific probabilities of finding each particle in a given region of space; photons, on the other hand, were viewed as nothing but excited states of the quantized electromagnetic field where photons can be freely created or destroyed. By analogy to the view that photons are excited states of the quantized electromagnetic field, each type of particle is represented within QFT by its corresponding quantum field. In short, QFT rests on two fundamental concepts:

- the fundamental entities of QFT are fields, not particles.
- Particles are viewed as excitations of fields; that is, a particle can be represented by the quantum oscillations of a field.

13.8 The Lagrangian of Quantum Electrodynamics

To illustrate the power of the Lagrangian formulation of field theories, we want to highlight few ideas about *Quantum Electrodynamics* (QED). In a nutshell, QED is the quantum counterpart of classical electrodynamics (Maxwell's theory). It deals with the interaction of light (i.e., photons) with charged particles, particularly electrons. In addition to describing the interactions of the electromagnetic field with matter, QED deals also with interactions between charged particles. QED is one of the most tested theories; its predictions agree with experiment to an accuracy of about 8 digits, making it one of the most accurate theories in physics.

[8] W. Pauli, *Rev. Mod. Phys.* **13**, 203 (1941).
C. N. Yang and R. L. Mills, *Phys. Rev.*, **96**, 191 (1954).
G.'t Hooft, *Nuclear Physics B*, **33**, 173 (1971); G.'t Hooft, *Nuclear Physics B*, **35**, 167 (1971).

In what follows, we are going to deal with few ideas about the Lagrangian formulation of QED. As mentioned above, Lagrangians are great theoretical tools for constructing field theories. Due to their scalar nature, Lagrangians are invariant under Lorentz transformations; they lend themselves quite readily to extensions into the relativistic domain. Additionally, Lagrangians handle symmetries with great ease. As an edifying illustration, we consider the interaction between a charged particle and an electromagnetic field. An electron which is coupled to an electromagnetic field is described by means of the Lagrangian density of Quantum Electrodynamics:

$$\mathcal{L}_{QED} = \mathcal{L}_{Dirac} + \mathcal{L}_{Int} + \mathcal{L}_{EMF}, \tag{13.114}$$

where $\mathcal{L}_{Dirac}$ accounts for the propagation of the free electron (see Eq. (13.88)):

$$\mathcal{L}_{Dirac}(\psi, \partial_\mu\psi, \overline{\psi}, \partial_\mu\overline{\psi}) = i\hbar c\overline{\psi}\gamma^\mu\partial_\mu\psi - mc^2\overline{\psi}\psi, \tag{13.115}$$

where $q = -e$ is the charge of the electron (note that e is positive); $\mathcal{L}_{Int}$ relates to the interaction between the electron and the electromagnetic field A^μ,

$$\mathcal{L}_{Int}(\psi, \overline{\psi}, A^\mu) = -qc\overline{\psi}\gamma_\mu\psi A^\mu; \tag{13.116}$$

and $\mathcal{L}_{EMF}$ governs the electromagnetic field in vacuum (see Eq. (13.52)):

$$\mathcal{L}_{EMF}(A_\mu, \partial^\mu A_\mu) = -\frac{1}{4\mu_0}F^{\mu\nu}F_{\mu\nu}, \tag{13.117}$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \equiv \frac{\partial A^\nu}{\partial x_\mu} - \frac{\partial A^\mu}{\partial x_\nu}$. Hence, we can rewrite the *QED Lagrangian density* (13.114) as follows

$$\boxed{\mathcal{L}_{QED}(\psi, \partial_\mu\psi, \overline{\psi}, \partial_\mu\overline{\psi}, A_\mu, \partial^\mu A_\mu) = \left(i\hbar c\overline{\psi}\gamma^\mu\partial_\mu\psi - mc^2\overline{\psi}\psi\right) - qc\overline{\psi}\gamma_\mu\psi A^\mu - \frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu}.} \tag{13.118}$$

An application of this Lagrangian density to the Euler–Lagrange equation yields both the Dirac equation and the Maxwell equations. As shown in Problem 13.7, to derive the Dirac equation (13.183), all we need to do is to vary the QED Lagrangian density (13.118) with respect to the Dirac spinor field $\overline{\psi}$:

$$\boxed{\frac{\partial\mathcal{L}_{QED}}{\partial\overline{\psi}} - \partial_\mu\left(\frac{\partial\mathcal{L}_{QED}}{\partial(\partial_\mu\overline{\psi})}\right) = 0 \quad \Longrightarrow \quad c\gamma_\mu\left(i\hbar\partial^\mu - qA^\mu\right)\psi - mc^2\psi = 0.} \tag{13.119}$$

Similarly, by varying (13.118) with respect to the electromagnetic field A^μ, we obtain easily the inhomogeneous Maxwell equation (13.62):

$$\boxed{\frac{\partial\mathcal{L}_{QED}}{\partial A_\nu} - \partial_\mu\left(\frac{\partial\mathcal{L}_{QED}}{\partial(\partial_\mu A_\mu)}\right) = 0 \quad \Longrightarrow \quad \partial_\mu F^{\mu\nu} = \mu_0 j^\nu.} \tag{13.120}$$

The above equations show that, with a *single* Lagrangian density, QED provides us with all the information we need to know about an electron or a photon when they are alone or when they are interacting with each other. For instance, in the absence of an electromagnetic field, and if we have an electron by itself, the QED Lagrangian density (13.118) reduces to

that of Dirac, $\mathcal{L}_{QED} = \mathcal{L}_{Dirac}$, and, hence, it can provide us with all what the Dirac equation is capable of offering. Additionally, in the absence of charged particles, and if we have an electromagnetic field that is propagating in space, the QED Lagrangian density reduces to that of an electromagnetic field, $\mathcal{L}_{QED} = \mathcal{L}_{EMF}$; thus, giving us back the Maxwell equations. In short, QED is equipped with the necessary conceptual as well as theoretical tools to deal with charged particles and their interactions with electromagnetic fields; it also deal with other phenomena such as pair production creation and electron-positron annihilation. We may also mention that QED is a mere subfield within QFT. As mentioned in Section 13.7; QFT is more general than QED. While QED deals with the elctromagnetic force only, QFT deals with the electromagnetic, weak, and strong forces.

13.9 Concluding Thoughts

To appreciate the power of quantum field theory (QFT), we may mention some edifying examples about its success in addressing the limitations of relativistic quantum mechanics (RQM) and classical field theory:

- While RQM is the theory of particles, QFT is the theory of quantum field operators which deals with both particles *and* fields; RQM deals primarily with single particle states, whereas QFT with multi-particle states involving infinite degrees of freedom.

- Unlike RQM which cannot deal with the creation and destruction of particles, QFT offers a natural theoretical framework where particles and antiparticles can be created or destroyed; for instance, while QFT can handle the transmutation of particles from one kind into another, RQM cannot. In typical reactions like muon decay, $\mu^- \longrightarrow e^- + \nu_\mu + \bar{\nu}_e$, QFT uses field operators in annihilating the initial muon and creating a final electron and, in the process, giving rise to a muon neutrino and an anti-electron neutrino to balance the reaction so as to satisfy the relevant laws of conservation.

- The problem of negative energy states that has plagued RQM does not occur in QFT where the field operators create and destroy both particles and antiparticles and both of these have positive energies.

- The problem of negative probability densities that has caused undue confusion in RQM does not occur in QFT. Being multiparticle, QFT deals with charge densities; probability densities occur only in single-particle theories such as RQM.

- QFT has no need for the conceptually controversial hole theory of Dirac. As mentioned above, since the problem of negative energies does not occur in QFT, the conceptually flawed hypothesis of an infinite sea of negative energy states has no role to play in QFT.

- Unlike the Klein-Gordon equation which deals with bosons only, and the Dirac equation which applies only to fermions, QFT has the theoretical tools to describe both bosonic as well as fermionic systems.

- As their names suggest, QFT is quantum mechanical, classical field theory is not. Unlike QFT, classical field theory is not equipped with the theoretical tools to offer solutions to the problems that have plagued RQM.

- In short, QFT, which unifies quantum mechanics and special relativity, forms the basis of our modern understanding of how the building blocks of nature interact.

Although quantum field theory was developed mainly for the field of elementary particle physics, it has been applied successfully to other branches of physics that deal with interacting many-particle systems, most notably to condensed matter physics and statistical mechanics. In the field of condensed matter physics, QFT has been applied with remarkable success to a wide range of problems such as organic conductors, carbon nanotubes, high-Tc superconductors, quantum Hall systems, and to antiferromagnetic chains, etc. In the area of statistical mechanics, QFT is especially useful in the study of quantum phase transitions and critical phenomena.

In summary, quantum field theory is the best theoretical framework available to us to describe interacting many-particle systems. Cutting across several areas of physics and due to its great quantitative predictive power, QFT has proven to be the most accurate theory in science; its various quantitative results have been verified experimentally with great accuracy, sometimes to about eight significant figures.

13.10 Solved Problems

Problem 13.1
Using the Hamiltonian formulation of classical fields, find the Hamiltonian for a free charged particle (of mass m and charge q) moving in an external electromagnetic field consisting of an electric scalar potential Φ and a vector potential $\vec{A}$.

Solution
First, the Hamiltonian of a *free* particle is given by

$$H = \frac{1}{2m}\vec{p}^{\,2}. \tag{13.121}$$

Now, when the particle interacts with an external electromagnetic fiedl, its classical Hamiltonian can be inferred from $H(q,\ p;\ t) = \sum_{j=1}^{n} \dot{q}_j p_j - L(q,\ \dot{q};\ t)$ which was derived in Appendix F:

$$H = \sum_{j=1}^{3} \dot{q}_j p_j - L = \dot{x}p_x + \dot{y}p_y + \dot{z}p_z - L = \dot{\vec{r}}\cdot\vec{p} - L \equiv \vec{v}\cdot\vec{p} - L, \tag{13.122}$$

where $\dot{\vec{r}} \equiv \vec{v}$ is the particle's velocity, $\vec{p}$ is the generalized momentum, and L is the Lagrangian, $L = T - V$. So, the problem reduces to finding the potential energy V which results from the interaction of the particle with the external electromagnetic field. From the classical theory of electromagnetism, when a charge q moves with a velocity $\vec{v}$ under the influence of an electric field $\vec{E}$ and magnetic field $\vec{B}$, the force of interaction between the particle and the electromagnetic field is given by the Lorentz force

$$\vec{F} = q\left(\vec{E} + \vec{v}\times\vec{B}\right). \tag{13.123}$$

The electric and magnetic fields are expressed in terms of the scalar and vector potentials, $U(\vec{r})$ and $\vec{A}(\vec{r})$, as follows:

$$\vec{E} = -\vec{\nabla}U - \frac{\partial \vec{A}}{\partial t}, \qquad \vec{B} = \vec{\nabla} \times \vec{A}. \tag{13.124}$$

Inserting (13.124) into (13.123), we can show that the Lorentz force becomes

$$\vec{F} = q\left[-\vec{\nabla}U - \frac{\partial \vec{A}}{\partial t} + \nabla(\vec{v} \cdot \vec{A})\right]. \tag{13.125}$$

For instance, we can write the $x-$component of the force as

$$F_x = q\left[-\frac{\partial U}{\partial x} - \frac{\partial A_x}{\partial t} + \frac{\partial}{\partial x}(\vec{v} \cdot \vec{A})\right], \tag{13.126}$$

or as

$$F_x = q\left[-\frac{\partial U}{\partial x} - \frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}}(\vec{v} \cdot \vec{A})\right) + \frac{\partial}{\partial x}(\vec{v} \cdot \vec{A})\right], \tag{13.127}$$

which, in turn, can be rewritten in terms of a scalar potential energy V as follows

$$F_x = -\frac{\partial V}{\partial x} + \frac{d}{dt}\left(\frac{\partial V}{\partial \dot{x}}\right), \tag{13.128}$$

where

$$V = qU - q\vec{v} \cdot \vec{A}. \tag{13.129}$$

Hence, the particle's Lagrangian is given by

$$L = T - V = T - qU + q\vec{v} \cdot \vec{A} \equiv \frac{1}{2m}\vec{v}^{\,2} - qU + q\vec{v} \cdot \vec{A}. \tag{13.130}$$

We now can obtain the generalized momentum for this charged particle; since the scalar potential U does not depend on $\dot{x}$, we can obtain, for instance, the $x-$component as follows

$$p_x = \frac{\partial L}{\partial \dot{x}} \equiv \frac{\partial}{\partial \dot{x}}\left(\frac{1}{2m}\dot{x}^2 - qU + q\dot{x}A_x\right) = m\dot{x} + qA_x \tag{13.131}$$

Hence, we can write the generalized momentum as

$$\vec{p} = m\vec{v} + q\vec{A}, \tag{13.132}$$

and, hence, obtain the particle's velocity

$$\vec{v} = \frac{1}{m}\left(\vec{p} - q\vec{A}\right). \tag{13.133}$$

Now, inserting (13.132) into (13.122), we obtain

$$\begin{aligned} H &= \vec{v} \cdot \vec{p} - L \\ &= \vec{v} \cdot \left(m\vec{v} + q\vec{A}\right) - L \\ &= m\vec{v}^{\,2} + q\vec{v} \cdot \vec{A} - L. \end{aligned} \tag{13.134}$$

A simple insertion of (13.130) into (13.134 leads to

$$\begin{aligned} H &= m\vec{v}^2 + q\vec{v}\cdot\vec{A} - L \\ &= m\vec{v}^2 + q\vec{v}\cdot\vec{A} - \frac{1}{2m}\vec{v}^2 + qU - q\vec{v}\cdot\vec{A} \\ &= \frac{1}{2}m\vec{v}^2 + qU. \end{aligned} \tag{13.135}$$

Finally, we can obtain the Hamiltonian by simply substituting (13.133) into (13.134:

$$H = \frac{1}{2}m\vec{v}^2 + qU = \frac{1}{2m}\left(\vec{p} - q\vec{A}\right)^2 + qU. \tag{13.136}$$

Remark
Comparing (13.121) and (13.136), it is easy to see that we can obtain the Hamiltonian of a particle of charge q in an electromagnetic field by simply replacing the following relations

$$\boxed{H_0 \longrightarrow H = H_0 - qU, \qquad \vec{p} \longrightarrow \vec{p} = \vec{p}_0 - q\vec{A},} \tag{13.137}$$

into the *free* particle Hamiltonian (13.121), $H - 0 = \vec{p}^2/2m$; that is,

$$\boxed{H_0 = \frac{1}{2m}\vec{p}^2 \longrightarrow H = H_0 - qU = \frac{1}{2m}\left(\vec{p}^2 - q\vec{A}\right)^2 \quad \Longrightarrow \quad H = \frac{1}{2m}\left(\vec{p} - q\vec{A}\right)^2 + qU.} \tag{13.138}$$

The relations (13.137) are known as the *minimal substitution* or *minimal coupling* principle. We have used them in Chapter 6 (see Eq. (6.217)) to derive the Pauli equation for a spin 1/2 particle in an electromagnetic field and in Chapter 12 (see Eq. (12.123)) to study the Dirac equation in an electromagnetic field. This procedure states that the Hamiltonian for a charged particle, q, that is interacting with an external electromagnetic field can be obtained by making the substitutions (13.137) into the *free* particle Hamiltonian.

Problem 13.2
If T_S is a symmetric second-rank tensor and T_A an antisymmetric tensor, show that $T_{S\mu\nu}T_A^{\mu\nu} = 0$.

Solution
On the one hand, since $T_{S\mu\nu} = T_{S\nu\mu}$ and $T_A^{\mu\nu} = -T_A^{\nu\mu}$, we can write

$$T_{S\mu\nu}T_A^{\mu\nu} = -T_{S\nu\mu}T_A^{\nu\mu}. \tag{13.139}$$

On the other hand, since the summation indices μ and ν are dummy variables, we can rename them to ν and μ, respectively, and, hence, write

$$T_{S\mu\nu}T_A^{\mu\nu} \equiv T_{S\nu\mu}T_A^{\nu\mu}. \tag{13.140}$$

A comparison of (13.139) and (13.140) shows that

$$T_{S\mu\nu}T_A^{\mu\nu} = -T_{S\nu\mu}T_A^{\nu\mu} = T_{S\nu\mu}T_A^{\nu\mu} \qquad \Longrightarrow \qquad T_{S\mu\nu}T_A^{\mu\nu} = 0, \tag{13.141}$$

since when a number is equal to its opposite, it must be zero.

Problem 13.3
(a) Using the electric and magnetic field relations $\vec{E} = -\nabla U - \frac{\partial \vec{A}}{\partial t}$ and $\vec{B} = \vec{\nabla} \times \vec{A}$, show that they lead to the electromagnetic field tensor $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$.
(b) Using the result of part (a), find the matrix representation of $F^{\mu\nu}$.

Solution

(a) In this problem, we are going to use the index notation of 4-vectors. We can write the 4-vector potential A^μ, where U is the electric scalar potential and $\vec{A}$ is the 3-vector potential, and the gradient operators ∂^μ and ∂_μ as follows

$$A^\mu = \left(U/c, \vec{A}\right) = (A^0, A^1, A^2, A^3), \tag{13.142}$$

$$\partial^\mu = \left(\partial^0, \partial^1, \partial^2, \partial^3\right) = \frac{\partial}{\partial x_\mu} \equiv \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right), \tag{13.143}$$

$$\partial_\mu = (\partial_0, \partial_1, \partial_2, \partial_3) = \frac{\partial}{\partial x^\mu} \equiv \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right), \tag{13.144}$$

where we have used the facts that $x^\mu = (ct, x, y, z) = (x^0, x^1, x^2, x^3)$ and $x_\mu = (ct, -x, -y, -z) = (x_0, x_1, x_2, x_3)$. We can also write the components of $\vec{E} = -\nabla U - \partial \vec{A}/\partial t$ as follows:

$$E_x = -\frac{\partial U}{\partial x} - \frac{\partial A_x}{\partial t} \equiv \frac{\partial (cA^0)}{\partial x_1} - c\frac{\partial A^1}{\partial x_0} \quad \Longrightarrow \quad E^1 = -c\left(\partial^0 A^1 - \partial^1 A^0\right), \tag{13.145}$$

$$E_y = -\frac{\partial U}{\partial y} - \frac{\partial A_y}{\partial t} \equiv \frac{\partial (cA^0)}{\partial x_2} - c\frac{\partial A^2}{\partial x_0} \quad \Longrightarrow \quad E^2 = -c\left(\partial^0 A^2 - \partial^2 A^0\right), \tag{13.146}$$

$$E_z = -\frac{\partial U}{\partial z} - \frac{\partial A_z}{\partial t} \equiv \frac{\partial (cA^0)}{\partial x_3} - c\frac{\partial A^3}{\partial x_0} \quad \Longrightarrow \quad E^3 = -c\left(\partial^0 A^3 - \partial^3 A^0\right), \tag{13.147}$$

Similarly, we can write the components of $\vec{B} = \vec{\nabla} \times \vec{A}$ as follows:

$$B_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \equiv -\frac{\partial A^3}{\partial x_2} + \frac{\partial A^2}{\partial x_3} \quad \Longrightarrow \quad B^1 = -\left(\partial^2 A^3 - \partial^3 A^2\right), \tag{13.148}$$

$$B_y = \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \equiv -\frac{\partial A^1}{\partial x_3} + \frac{\partial A^3}{\partial x_1} \quad \Longrightarrow \quad B^2 = -\left(\partial^3 A^1 - \partial^1 A^3\right), \tag{13.149}$$

$$B_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \equiv -\frac{\partial A^2}{\partial x_1} + \frac{\partial A^1}{\partial x_2} \quad \Longrightarrow \quad B^3 = -\left(\partial^1 A^2 - \partial^2 A^1\right), \tag{13.150}$$

From the six relations above, we see that the components of the electric and magnetic fields are connected; this is expected, since the electric and magnetic fields of an electromagnetic field are related. In relativistic notation, the six components E_x, E_y, E_z, B_x, B_y, B_z can be assembled together to form a second-rank tensor which we call the electromagnetic field strength tensor, $F^{\mu\nu}$. Each of these six components can be set to one component of $F^{\mu\nu}$. For instance, the structures of the relations (13.145)–(13.147) suggest to write the components F^{0i} as follows:

$$F^{01} = \partial^0 A^1 - \partial^1 A^0 = -E_x/c, \qquad F^{02} = \partial^0 A^2 - \partial^2 A^= - E_y/c, \qquad F^{03} = \partial^0 A^3 - \partial^3 A^0 = -E_z/c. \tag{13.151}$$

Similarly, equations (13.148)–(13.150) offer sufficient clues to write some of the components F^{ij} as follows: and

$$F^{23} = \partial^2 A^3 - \partial^3 A^2 = -B_x, \qquad F^{13} = \partial^1 A^3 - \partial^3 A^1 = B_y, \qquad F^{12} = \partial^1 A^2 - \partial^2 A^1 = -B_z. \tag{13.152}$$

Equations (13.151) and (13.152) can be condensed, respectively, as follows:

$$F^{0i} = \partial^0 A^i - \partial^i A^0 = -E_i/c, \qquad i = 1, 2, 3, \tag{13.153}$$

$$F^{ij} = \partial^i A^j - \partial^j A^i = \epsilon_{ijk} B_k, \qquad i, j, k = 1, 2, 3, \tag{13.154}$$

where ϵ_{ijk} is the 3-dimensional *Levi-Civita tensor*:

$$\epsilon_{ijk} = \begin{cases} -1 & \text{if } (i, j, k) \text{ is an } \textit{odd} \text{ permutation of } (1,2,3) \\ +1 & \text{if } (i, j, k) \text{ is an } \textit{even} \text{ permutation of } (1,2,3) \\ 0 & \text{for all other cases} \end{cases} \tag{13.155}$$

Finally, by combining (13.153) and (13.154), we can infer the form of $F^{\mu\nu}$:

$$\boxed{F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \equiv \frac{\partial A^\nu}{\partial x_\mu} - \frac{\partial A^\mu}{\partial x_\nu}.} \tag{13.156}$$

(b) The matrix representation of the electromagnetic field tensor $F^{\mu\nu}$ can be easily inferred from (13.153), (13.154) and (13.156). First, from (13.156), it is easy to see that the 4 diagonal elements are zero; namely, $F^{\mu\mu} = \partial^\mu A^\mu - \partial^\mu A^\mu = 0$. Second, (13.156) implies that $F^{\mu\nu}$ is antisymmetric, $F^{\mu\nu} = -F^{\nu\mu}$. Hence, we can obtain the remaining 6 sub-diagonal elements of $F^{\mu\nu}$ from (13.151)–(13.152) by using the antisymmetry property. In summary, putting all these conclusions together, we can easily infer the following matrix form of $F^{\mu\nu}$:

$$\boxed{(F^{\mu\nu}) = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}.} \tag{13.157}$$

Problem 13.4
Show how to obtain the inhomogeneous Maxwell equations (13.16)–(13.19) from $\partial_\mu F^{\mu\nu} = \mu_0 j^\nu$.

Solution
To obtain the first inhomogeneous Maxwell equation, $\vec{\nabla} \cdot \vec{E} = \rho/\epsilon_0$, we need only to apply ∂_μ on $F^{\mu 0}$. To see this, and using the relations $F^{00} = 0$ and $F^{i0} = E_i/c$, we can write

$$\begin{aligned} \partial_\mu F^{\mu 0} &= \partial_0 F^{00} + \partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30} = \partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30} \\ &= \frac{1}{c}\left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}\right) \\ &= \frac{1}{c}\vec{\nabla} \cdot \vec{E}. \end{aligned} \tag{13.158}$$

Next, by inserting this result along with $j^0 = c\rho$ into $\partial_\mu F^{\mu 0} = \mu_0 j^0$, we obtain

$$\partial_\mu F^{\mu 0} = \mu_0 j^0 \quad \Longrightarrow \quad \frac{1}{c}\vec{\nabla}\cdot\vec{E} = \mu_0 c\rho \quad \Longrightarrow \quad \vec{\nabla}\cdot\vec{E} = \mu_0 c^2 \rho \quad \Longrightarrow \quad \vec{\nabla}\cdot\vec{E} = \frac{\rho}{\varepsilon_0}, \tag{13.159}$$

where we have used the fact relation $c^2 = 1/(\varepsilon_0\mu_0)$ in the last step.

For the second inhomogeneous Maxwell equation, $\vec{\nabla}\times\vec{B} - \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t} = \mu_0\vec{J}$, we simpl need to apply ∂_μ on $F^{\mu i}$. For instance, taking $i = 1$, and using the relations $F^{01} = -E_x/c$, $F^{11} = 0$, $F^{21} = B_z$, $F^{31} = -B_y$, we can write

$$\begin{aligned}
\partial_\mu F^{\mu 1} &= \partial_0 F^{01} + \partial_1 F^{11} + \partial_2 F^{21} + \partial_3 F^{31} = \partial_0 F^{01} + \left(\partial_2 F^{21} + \partial_3 F^{31}\right) \\
&= -\frac{1}{c}\frac{\partial E_x}{\partial(ct)} + \left(\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z}\right) = \left(\vec{\nabla}\times\vec{B}\right)_x - \frac{1}{c^2}\left(\frac{\partial\vec{E}}{\partial t}\right)_x \\
&= \left[\left(\vec{\nabla}\times\vec{B}\right) - \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t}\right]_x .
\end{aligned} \tag{13.160}$$

Inserting this result into $\partial_\mu F^{\mu 1} = \mu_0 j^1$ and since $j^1 = J_x$, we obtain

$$\partial_\mu F^{\mu 1} = \mu_0 j^1 \quad \Longrightarrow \quad \left[\left(\vec{\nabla}\times\vec{B}\right) - \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t}\right]_x = \mu_0 J_x. \tag{13.161}$$

Following the same method by applying ∂_μ on $F^{\mu 2}$ and on $F^{\mu 3}$, we obtain

$$\partial_\mu F^{\mu 2} = \mu_0 j^2 \quad \Longrightarrow \quad \left[\left(\vec{\nabla}\times\vec{B}\right) - \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t}\right]_y = \mu_0 J_y, \tag{13.162}$$

$$\partial_\mu F^{\mu 3} = \mu_0 j^3 \quad \Longrightarrow \quad \left[\left(\vec{\nabla}\times\vec{B}\right) - \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t}\right]_z = \mu_0 J_z. \tag{13.163}$$

Finally, putting (13.161)–(13.163) together, we obtain:

$$\left(\vec{\nabla}\times\vec{B}\right) - \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t} = \mu_0\vec{J} \quad \text{or} \quad \left(\vec{\nabla}\times\vec{B}\right) - \varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t} = \mu_0\vec{J} \tag{13.164}$$

This is the second Maxwell equation.

For the third (homogeneous) Maxwell equation, $\vec{\nabla}\cdot\vec{B} = 0$, we proceed as follows:

$$\vec{\nabla}\cdot\vec{B} = 0 \quad \Longrightarrow \quad \partial_1 F^{32} + \partial_2 F^{13} + \partial_3 F^{21} = 0, \tag{13.165}$$

where we have used the relations: $F^{32} = B_x$, $F^{13} = B_y$, $F^{21} = B_z$.

Finally, for the fourth (homogeneous) Maxwell equation, we have

$$\left[\vec{\nabla}\times\vec{E} + \frac{\partial\vec{B}}{\partial t}\right]_x = 0 \quad \Longrightarrow \quad \partial_2 F^{30} + \partial_3 F^{02} + \partial_0 F^{32} = 0, \tag{13.166}$$

$$\left[\vec{\nabla}\times\vec{E}+\frac{\partial\vec{B}}{\partial t}\right]_y=0 \quad\Longrightarrow\quad \partial_3F^{10}+\partial_1F^{03}+\partial_0F^{13}=0, \tag{13.167}$$

$$\left[\vec{\nabla}\times\vec{E}+\frac{\partial\vec{B}}{\partial t}\right]_z=0 \quad\Longrightarrow\quad \partial_1F^{20}+\partial_2F^{01}+\partial_0F^{21}=0, \tag{13.168}$$

We can combine the previous four relations (i.e., the two homogeneous Maxwell equations) into a single equation:

$$\boxed{\partial_\lambda F^{\mu\nu}+\partial_\mu F^{\nu\lambda}+\partial_\nu F^{\lambda\mu}=0 \quad\Longrightarrow\quad \tfrac{1}{2}\epsilon^{\lambda\mu\nu\alpha}F_{\nu\alpha}=0,} \tag{13.169}$$

where $\lambda,\ \mu,\ \nu,\ \alpha = 0,\ 1,\ 2,\ 3$ and where $\epsilon^{\lambda\mu\nu\alpha}$ is the 4-dimensional *Levi-Civita tensor*:

$$\epsilon^{\lambda\mu\nu\alpha}=\begin{cases}-1 & \text{if } (\lambda,\ \mu,\ \nu,\ \alpha) \text{ is an } \textit{odd} \text{ permutation of } (0,\ 1,\ 2,\ 3)\\ +1 & \text{if } (\lambda,\ \mu,\ \nu,\ \alpha) \text{ is an } \textit{even} \text{ permutation of } (0,\ 1,\ 2,\ 3)\\ 0 & \text{for all other cases}\end{cases} \tag{13.170}$$

Equation (13.169) is known as the *Bianchi identity* for the field strength tensor $F^{\mu\nu}$.

In summary, while the two inhomogeneous Maxwell equations, (13.16) and (13.17), can be obtained from $\partial_\mu F^{\mu\nu}=\mu_0 j^\nu$, the two homogeneous equations, (13.18) and (13.19), can be expressed by $\partial_\lambda F^{\mu\nu}+\partial_\mu F^{\nu\lambda}+\partial_\nu F^{\lambda\mu}=0$.

Problem 13.5

Consider the following Lagrangian density which describes the propagation of a massive spin 1 (vector) field A^μ in free space:

$$\mathcal{L}=-\frac{1}{4\mu_0}\left(\partial_\mu A_\nu-\partial_\nu A_\mu\right)\left(\partial^\mu A^\nu-\partial^\nu A^\mu\right)+\frac{1}{2\mu_0}\left(\frac{mc}{\hbar}\right)^2 A_\mu A^\mu.$$

(a) Derive the equations of motion for the field A^μ. Is this equation familiar?
(b) Show that the equations derived in (a) reduce to the Maxwell equations for $m=0$.
(c) Show that by applying ∂_ν to the equation derived in (a), we obtain $\partial_\nu A^\nu=0$.
(d) Show that if $\partial_\nu A^\nu=0$, the equation derived in (a) reduce to the Klein-Gordon equation.

Solution

(a) First, using the relation $F^{\mu\nu}=\partial^\mu A^\nu-\partial^\nu A^\mu$, we can write the given Lagrangian density in the following simple form

$$\mathcal{L}=-\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu}+\frac{1}{2\mu_0}\left(\frac{mc}{\hbar}\right)^2 A_\nu A^\nu. \tag{13.171}$$

Note that the above relation is similar to the Maxwell Lagrangian density to which we have added a "massive" term $\frac{1}{2\mu_0}\left(\frac{mc}{\hbar}\right)^2 A_\nu A^\nu$.

We need to apply (13.171) to the Euler–Lagrange equation

$$\frac{\partial\mathcal{L}}{\partial A_\nu}-\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu A_\nu\right)}\right)=0. \tag{13.172}$$

First, using (13.171), we have

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial A_\nu} &= \frac{\partial}{\partial A_\nu}\left[-\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu}+\frac{1}{2\mu_0}\left(\frac{mc}{\hbar}\right)^2 A_\nu A^\nu\right] \\ &= \frac{\partial}{\partial A_\nu}\left[\frac{1}{2\mu_0}\left(\frac{mc}{\hbar}\right)^2 A_\nu A^\nu\right] \\ &= \frac{1}{\mu_0}\left(\frac{mc}{\hbar}\right)^2 A^\nu. \end{aligned} \tag{13.173}$$

Second, using the expression $F_{\mu\nu}F^{\mu\nu} = 2\left(\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu\right) - 2\left(\partial_\mu A_\nu\right)\left(\partial^\nu A^\mu\right)$ that is listed in (13.53), we can obtain the following derivative

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu A_\nu\right)} &= \frac{\partial}{\partial\left(\partial^\mu A^\nu\right)}\left[-\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu}+\frac{1}{2\mu_0}\left(\frac{mc}{\hbar}\right)^2 A^\mu A_\mu\right] \\ &= -\frac{1}{4\mu_0}\frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left(F_{\mu\nu}F^{\mu\nu}\right) \\ &= -\frac{1}{2\mu_0}\frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left[\left(\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu\right) - \left(\partial_\mu A_\nu\right)\left(\partial^\nu A^\mu\right)\right] \\ &= -\frac{1}{\mu_0}\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) \\ &= -\frac{1}{\mu_0}F^{\mu\nu}. \end{aligned} \tag{13.174}$$

Finally, inserting (13.173) and (13.174) into (13.172), we obtain

$$\boxed{\partial_\mu F^{\mu\nu} + \left(\tfrac{mc}{\hbar}\right)^2 A^\nu = 0.} \tag{13.175}$$

This is known as the *Proca equation*[9]. It is used to describe the propagation of a massive spin-1 field in free space where no source of charges or currents are present.

(b) If $m = 0$, the Proca relation (13.175) reduces to the homogeneous Maxwell equations (see Eq. (13.58)): $\partial_\mu F^{\mu\nu} = 0$.

(c) By applying ion ∂_ν to (13.175), we obtain

$$\partial_\nu\partial_\mu F^{\mu\nu} + \left(\frac{mc}{\hbar}\right)^2 \partial_\nu A^\nu = 0 \qquad \Longrightarrow \qquad \partial_\nu A^\nu = 0, \tag{13.176}$$

since $\partial_\nu\partial_\mu F^{\mu\nu} = 0$ as demonstrated in (13.49).

(d) Using the relation $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ along with the condition $\partial_\mu A^\mu = 0$, we can write

$$\begin{aligned}\partial_\mu F^{\mu\nu} &= \partial_\mu\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) = \left(\partial_\mu\partial^\mu\right)A^\nu - \partial^\nu\left(\partial_\mu A^\mu\right) \\ &= \left(\partial_\mu\partial^\mu\right)A^\nu \\ &= \Box A^\nu, \end{aligned} \tag{13.177}$$

[9] A. Proca, *Compt. Rend.* **190**, 1377 (1930); *Compt. Rend.* **191**, 26 (1930); *J. Phys. Radium Ser. VII*, **1**, 235 (1936); *J. Phys. Radium Ser. VII*, **7**, 347 (1936); *J. Phys. Radium Ser. VII*, **8**, 23 (1937).

where we have used in the third line the d'Alembertian relation $\partial_\mu\partial^\mu = \Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$. Finally, by inserting (13.177) into (13.175), we see that the Proca relation (13.175) reduces to the following

$$\boxed{\partial_\mu F^{\mu\nu} + \left(\frac{mc}{\hbar}\right)^2 A^\nu = 0 \quad \text{and if} \quad \partial_\mu A^\mu = 0 \qquad \Longrightarrow \qquad \left(\Box + \frac{m^2c^2}{\hbar^2}\right)A^\nu = 0.} \tag{13.178}$$

This is the Klein-Gordon equation for the field A^ν. This relation shows that all four components of the Proca field A^ν satisfy the Klein-Gordon equation.

Problem 13.6
Using the Lagrangian formulation, derive the Dirac equation for a spin 1/2 particle of charge q interacting with an external electromagnetic field A^μ.

Solution

First, we need to construct the Lagrangian density for the spin 1/2 charged particle in an electromagnetic field. For this, we need to use the minimal substitution $i\hbar\partial^\mu \longrightarrow i\hbar\partial^\mu - qA^\mu$ in the Dirac Lagrangian density (13.88):

$$\mathcal{L} = i\hbar c\overline{\psi}\gamma_\mu\partial^\mu\psi - mc^2\overline{\psi}\psi \quad \longrightarrow \quad \mathcal{L} = c\overline{\psi}\gamma_\mu\left(i\hbar\partial^\mu - qA^\mu\right)\psi - mc^2\overline{\psi}\psi. \tag{13.179}$$

Using this Lagrangian density and applying the Euler–Lagrange equation

$$\frac{\partial\mathcal{L}}{\partial\overline{\psi}} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\overline{\psi}\right)}\right) = 0, \tag{13.180}$$

to $\overline{\psi}$, we obtain

$$\frac{\partial\mathcal{L}}{\partial\overline{\psi}} = \frac{\partial}{\partial\overline{\psi}}\left[c\overline{\psi}\gamma_\mu\left(i\hbar\partial^\mu - qA^\mu\right)\psi - mc^2\overline{\psi}\psi\right] = c\gamma_\mu\left(i\hbar\partial^\mu - qA^\mu\right)\psi - mc^2\psi, \tag{13.181}$$

and

$$\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\overline{\psi}\right)} = \frac{\partial\left[c\overline{\psi}\gamma_\mu\left(i\hbar\partial^\mu - qA^\mu\right)\psi - mc^2\overline{\psi}\psi\right]}{\partial\left(\partial_\mu\overline{\psi}\right)} = 0, \tag{13.182}$$

since $\mathcal{L}$ does not depend on $\partial_\mu\overline{\psi}$.

Now, inserting (13.181) and (13.182) into (13.180), we obtain

$$\boxed{c\gamma_\mu\left(i\hbar\partial^\mu - qA^\mu\right)\psi - mc^2\psi = 0.} \tag{13.183}$$

This is the covariant form of the Dirac equation for a a spin 1/2 particle of charge q interacting with an external electromagnetic field A^μ.

In what follows, we are going to reduce (13.183) to a standard differential form. For this, we need to find the expressions of $\gamma_\mu\partial^\mu$ and $\gamma_\mu A^\mu$. Using the inner product relation $B^\mu C_\mu = B^0C^0 - \vec{B}\cdot\vec{C}$ between any two 4-vectors $B^\mu = (B^0, B^1, B^2, B^3)$ and $C^\mu = (C^0, C^1, C^2, C^3)$

and since $A^\mu = (U/c, \vec{A})$, where U is the electric scalar potential and $\vec{A}$ is the vector potential, and since $\partial^\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\vec{\nabla}\right)$ and $\gamma_\mu = (\gamma^0, -\vec{\gamma}) \equiv (\beta, -\beta\vec{\alpha})$, we can write:

$$\gamma_\mu \partial^\mu = \gamma^0 \partial^0 - \vec{\gamma} \cdot \vec{\nabla} = \frac{1}{c}\beta\frac{\partial}{\partial t} - \beta\vec{\alpha} \cdot \vec{\nabla}, \tag{13.184}$$

and

$$\gamma_\mu A^\mu = \gamma^0 A^0 + \vec{\gamma} \cdot \vec{A} = \beta\frac{U}{c} + \beta\vec{\alpha} \cdot \vec{A}, \tag{13.185}$$

Multiplying (13.184) by $i\hbar$ and (13.185) by q and subtracting the second from the first, we obtain

$$\begin{aligned} i\hbar\gamma_\mu\partial^\mu - q\gamma_\mu A^\mu &= \frac{i\hbar}{c}\beta\frac{\partial}{\partial t} - i\hbar\beta\vec{\alpha} \cdot \vec{\nabla} - q\left(\beta\frac{U}{c} + \beta\vec{\alpha} \cdot \vec{A}\right) \\ &= \frac{i\hbar}{c}\beta\frac{\partial}{\partial t} + \beta\vec{\alpha} \cdot \left(-i\hbar\vec{\nabla} - q\vec{A}\right) - \frac{qU}{c}\beta, \end{aligned} \tag{13.186}$$

which, when multiplied by c, leads to

$$c\left(i\hbar\gamma_\mu\partial^\mu - q\gamma_\mu A^\mu\right) = i\hbar\beta\frac{\partial}{\partial t} + c\beta\vec{\alpha} \cdot \left(-i\hbar\vec{\nabla} - q\vec{A}\right) - qU\beta, \tag{13.187}$$

Now, inserting (13.187) into (13.183), we obtain

$$i\hbar\beta\frac{\partial\psi}{\partial t} - c\beta\vec{\alpha} \cdot \left(-i\hbar\vec{\nabla} - q\vec{A}\right)\psi - qU\beta\psi - mc^2\psi = 0. \tag{13.188}$$

After multiplying this equation by β and since $\beta^2 = 1$ and then rearranging the various terms, we obtain

$$i\hbar\frac{\partial\psi}{\partial t} = c\vec{\alpha} \cdot \left(-i\hbar\vec{\nabla} - q\vec{A}\right)\psi + qU\psi + mc^2\beta\psi, \tag{13.189}$$

or

$$i\hbar\frac{\partial\psi}{\partial t} = \left[c\vec{\alpha} \cdot \left(\hat{\vec{p}} - q\vec{A}\right) + mc^2\beta + \hat{V}\right]\psi, \tag{13.190}$$

where $\hat{\vec{p}} = -i\hbar\vec{\nabla}$ is the momentum operator and where we have used the notation $\hat{V} = qU$. This is the Dirac equation for a charged particle q in an electromagnetic field which we have derived in Chapter 12 (see Eq. (12.454)).

Problem 13.7

Consider an electron that is coupled to an electromagnetic field; i.e., an electron and a photon (massless vector boson of spin 1).

(a) Using the QED Lagrangian density (13.118), show how to obtain the Dirac equation for the electron (when interacting with the electromagnetic field A^μ).

(b) Using the QED Lagrangian density (13.118), show how to obtain the equation of the photon as it propagates in empty space.

Solution

(a) To obtain the equation of the electron, we need simply to use the Euler–Lagrange equation for the Dirac spinor field $\overline{\psi}$:

$$\frac{\partial \mathcal{L}_{QED}}{\partial \overline{\psi}} - \partial_\mu \left(\frac{\partial \mathcal{L}_{QED}}{\partial \left(\partial_\mu \overline{\psi} \right)} \right) = 0. \tag{13.191}$$

An application of the above equation to the QED Lagrangian density (13.118) allows us to calculate the first term as follows

$$\begin{aligned} \frac{\partial \mathcal{L}_{QED}}{\partial \overline{\psi}} &= \frac{\partial}{\partial \overline{\psi}} \left[c\overline{\psi}\gamma_\mu \left(i\hbar\partial^\mu - qA^\mu \right) \psi - mc^2 \overline{\psi}\psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right] \\ &= \frac{\partial}{\partial \overline{\psi}} \left[c\overline{\psi}\gamma_\mu \left(i\hbar\partial^\mu - qA^\mu \right) \psi - mc^2 \overline{\psi}\psi \right] \\ &= = c\gamma_\mu \left(i\hbar\partial^\mu - qA^\mu \right) \psi - mc^2 \psi, \end{aligned} \tag{13.192}$$

where we have used the fact that $F_{\mu\nu}F^{\mu\nu}$ does not depend on $\overline{\psi}$. As for the second term of (13.191), we have:

$$\frac{\partial \mathcal{L}_{QED}}{\partial \left(\partial_\mu \overline{\psi} \right)} = \frac{\partial}{\partial \left(\partial_\mu \overline{\psi} \right)} \left[c\overline{\psi}\gamma_\mu \left(i\hbar\partial^\mu - qA^\mu \right) \psi - mc^2 \overline{\psi}\psi - \frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} \right] = 0, \tag{13.193}$$

where we have used the fact that $\mathcal{L}_{QED}$ does not depend on $\partial_\mu \overline{\psi}$. Next, substituting (13.192) and (13.193) into (13.191), we obtain

$$\boxed{c\gamma_\mu \left(i\hbar\partial^\mu - qA^\mu \right) \psi - mc^2 \psi = 0.} \tag{13.194}$$

This is the Dirac equation we have derived above (see (13.183)).

(b) The Euler–Lagrange equation pertaining to the electromagnetic field, A^μ, is given by

$$\frac{\partial \mathcal{L}_{QED}}{\partial A_\nu} - \partial_\mu \left(\frac{\partial \mathcal{L}_{QED}}{\partial \left(\partial_\mu A_\nu \right)} \right) = 0. \tag{13.195}$$

Since $\mathcal{L}_{QED}$ depends on A_ν only through the term $\left(-cq\overline{\psi}\gamma_\nu \psi A^\mu \right)$, an applying the above relation to (13.118) leads to

$$\begin{aligned} \frac{\partial \mathcal{L}_{QED}}{\partial A_\nu} &= \frac{\partial}{\partial A^\nu} \left[c\overline{\psi}\gamma^\nu \left(i\hbar\partial_\nu - qA_\nu \right) \psi - mc^2 \overline{\psi}\psi - \frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} \right] \\ &= \frac{\partial}{\partial A_\nu} \left(-cq\overline{\psi}\gamma^\nu \psi A_\mu \right) \\ &= -cq\overline{\psi}\gamma^\nu \psi. \end{aligned} \tag{13.196}$$

Also, since $\mathcal{L}_{QED}$ depends on $\partial_\mu A_\nu$ only through the term $F^{\mu\nu}F_{\mu\nu}$, we can write

$$\begin{aligned}
\frac{\partial \mathcal{L}_{QED}}{\partial\left(\partial_\mu A_\nu\right)} &= \frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left[c\overline{\psi}\gamma^\nu\left(i\hbar\partial_\nu - qA_\nu\right)\psi - mc^2\overline{\psi}\psi - \frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu}\right] \\
&= -\frac{1}{4\mu_0}\frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left(F_{\mu\nu}F^{\mu\nu}\right) \\
&= -\frac{1}{2\mu_0}\frac{\partial}{\partial\left(\partial_\mu A_\nu\right)}\left[\left(\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu\right) - \left(\partial_\mu A_\nu\right)\left(\partial^\nu A^\mu\right)\right] \\
&= -\frac{1}{\mu_0}\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) \\
&= -\frac{1}{\mu_0}F^{\mu\nu}, \qquad (13.197)
\end{aligned}$$

where, in the third line, we have used the expression $F_{\mu\nu}F^{\mu\nu} = 2\left(\partial_\mu A_\nu\right)\left(\partial^\mu A^\nu\right) - 2\left(\partial_\mu A_\nu\right)\left(\partial^\nu A^\mu\right)$ that is listed in (13.53).

Finally, inserting (13.197) and (13.196) into (13.195), we obtain the equation for the photon

$$-cq\overline{\psi}\gamma^\nu\psi + \frac{1}{\mu_0}\partial_\mu F^{\mu\nu} = 0, \qquad (13.198)$$

which can be reduced to

$$\boxed{\partial_\mu F^{\mu\nu} = \mu_0 j^\nu,} \qquad (13.199)$$

where $j^\nu = cq\overline{\psi}\gamma^\nu\psi$ is the 4-current density. This is the inhomogeneous Maxwell equation in covariant form that we have derived in (13.62).

13.11 Exercises

Exercise 13.1

(a) If $T_S^{\mu\nu}$ is a symmetric second-rank tensor, how may independent elements does it have?

(b) If $T_A^{\mu\nu}$ is an antisymmetric second-rank tensor, how may independent elements does it have?

Exercise 13.2

Consider a general second-rank tensor $T^{\mu\nu}$ out of which we construct a symmetric tensor, $T_S{}^{\mu\nu} = \left(T^{\mu\nu} + T^{\nu\mu}\right)/2$ and an antisymmetric tensor, $T_A{}^{\mu\nu} = \left(T^{\mu\nu} - T^{\nu\mu}\right)/2$. Show that $T_{S\mu\nu}T_A^{\mu\nu} = 0$.

Exercise 13.3

Show that if $T_S^{\mu\nu}$ is a symmetric second-rank tensor, its Lorentz transform is also symmetric.

Exercise 13.4

Show that if $T_A^{\mu\nu}$ is an antisymmetric second-rank tensor, its Lorentz transform is also antisymmetric.

Exercise 13.5
Using the matrix form of the field tensor $F^{\mu\nu}$ listed in (13.35) and (13.36), show that $(F^{\mu\nu})(F_{\mu\nu})$ is a symmetric tensor.

Exercise 13.6
Consider the following Lagrangian density which describes the propagation of a massive spin 1 (vector) field A^μ in the presence of a source j^μ (4-vector current density):

$$\mathcal{L} = -\frac{1}{4\mu_0}\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right)\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) + \frac{1}{2}\left(\frac{mc}{\hbar}\right)^2 A^\mu A_\mu - j_\mu A^\mu,$$

(a) Derive the equations of motion for the field A^ν.

(b) Show that the equation derived in (a) reduces to the inhomogeneous Maxwell equations for $m = 0$.

(c) Show that if the condition $\partial_\nu A^\nu = 0$ is satisfied, the equation derived in (a) leads to the continuity condition $\partial_\nu j^\nu = 0$.

Exercise 13.7
Derive the equations of motion for the field A_μ from the following Lagrangian density

$$\mathcal{L} = -\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2\mu_0}\left(\frac{mc}{\hbar}\right)^2 A_\nu A^\nu - A_\nu j^\nu,$$

and show that, in free space, they reduce to the Klein-Gordon equations for the field A^μ when the condition $\partial_\mu A^\mu = 0$ is satisfied.

Exercise 13.8
Derive the equations of motion for the real scalar field ϕ from the following Lagrangian density

$$\mathcal{L}(\phi,\ \partial_\mu\phi) = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2c^2}{2\hbar^2}\phi^2 - \frac{\lambda}{4}\phi^4,$$

where λ is a coupling constant.

Appendix A

The Delta Function

A.1 One-Dimensional Delta Function

A.1.1 Various Definitions of the Delta Function

The delta function can be defined as the limit of $\delta^{(\varepsilon)}(x)$ when $\varepsilon \to 0$ (Figure A.1):

$$\delta(x) = \lim_{\varepsilon \to 0} \delta^{(\varepsilon)}(x), \tag{A.1}$$

where

$$\delta^{(\varepsilon)}(x) = \begin{cases} 1/\varepsilon, & -\varepsilon/2 < x < \varepsilon/2, \\ 0, & |x| > \varepsilon/2. \end{cases} \tag{A.2}$$

The delta function can be defined also by means of the following integral equations:

$$\int_{-\infty}^{+\infty} f(x)\delta(x)\,dx = f(0), \tag{A.3}$$

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a)\,dx = f(a). \tag{A.4}$$

We should mention that the δ-function is not a function in the usual mathematical sense. It can be expressed as the limit of analytical functions such as

$$\delta(x) = \lim_{\varepsilon \to 0} \frac{\sin(x/\varepsilon)}{\pi x}, \qquad \delta(x) = \lim_{a \to \infty} \frac{\sin^2(ax)}{\pi a x^2}, \tag{A.5}$$

or

$$\delta(x) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \frac{\varepsilon}{x^2 + \varepsilon^2}. \tag{A.6}$$

The Fourier transform of $\delta(x)$, which can be obtained from the limit of $\frac{\sin(x/\varepsilon)}{\pi x}$, is

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ikx} dk, \tag{A.7}$$

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

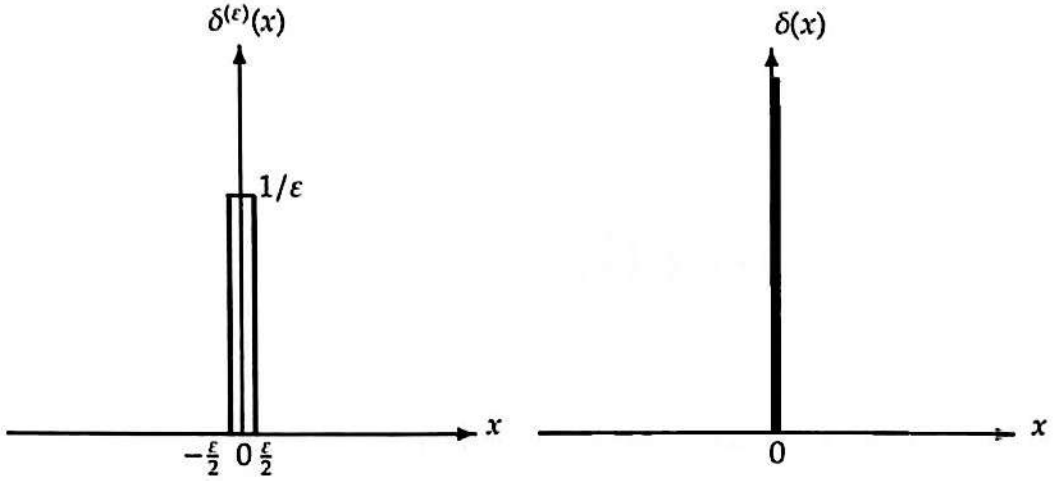


Figure A.1: The delta function $\delta(x)$ as defined by $\delta(x) = \lim_{\epsilon\to 0}\delta^{(\epsilon)}(x)$.

which in turn is equivalent to

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ikx}dk = \frac{1}{2\pi}\lim_{\epsilon\to 0}\int_{-1/\epsilon}^{+1/\epsilon} e^{ikx}dk = \lim_{\epsilon\to 0}\frac{\sin(x/\epsilon)}{\pi x} = \delta(x). \tag{A.8}$$

A.1.2 Properties of the Delta Function

The delta function is even:

$$\delta(-x) = \delta(x) \qquad \text{and} \qquad \delta(x-a) = \delta(a-x). \tag{A.9}$$

Here are some of the most useful properties of the delta function:

$$\int_a^b f(x)\delta(x-x_0)\,dx = \begin{cases} f(x_0), & \text{if } a < x_0 < b, \\ 0, & \text{elsewhere,} \end{cases} \tag{A.10}$$

$$\delta(x) = 0 \quad \text{for} \quad x \neq 0, \tag{A.11}$$

$$x\delta(x) = 0, \tag{A.12}$$

$$\delta(ax) = \frac{1}{|a|}\delta(x) \qquad (a \neq 0), \tag{A.13}$$

$$f(x)\delta(x-a) = f(a)\delta(x-a), \tag{A.14}$$

$$\int_c^d \delta(a-x)\delta(x-b)\,dx = \delta(a-b) \quad \text{for} \quad c \le a \le d, \quad c \le b \le d, \tag{A.15}$$

$$\int_a^b \delta(x)\,dx = 1 \quad \text{for} \quad a \le 0 \le b \tag{A.16}$$

$$\delta[g(x)] = \sum_i \frac{1}{|g'(x_i)|}\delta(x-x_i), \tag{A.17}$$

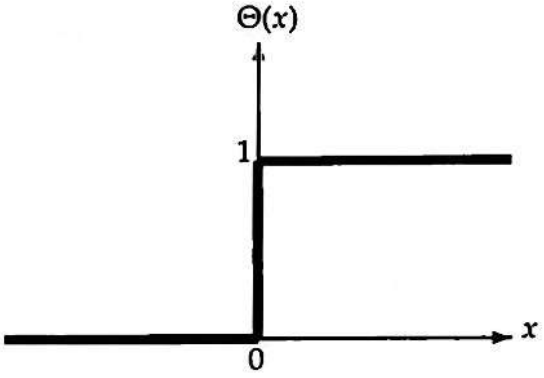


Figure A.2: The Heaviside function $\Theta(x)$.

where x_i is a zero of $g(x)$ and $g'(x_i) \neq 0$. Using (A.17), we can verify that

$$\delta[(x-a)(x-b)] = \frac{1}{|a-b|}[\delta(x-a)+\delta(x-b)] \qquad (a \neq b), \tag{A.18}$$

$$\delta(x^2-a^2) = \frac{1}{2|a|}[\delta(x-a)+\delta(x+a)] \qquad (a \neq 0). \tag{A.19}$$

A.1.3 Derivative of the Delta Function

The Heaviside function, or step function is defined as follows; see Figure A.2:

$$\Theta(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases} \tag{A.20}$$

The derivative of the Heaviside function gives back the delta function:

$$\frac{d}{dx}\Theta(x) = \delta(x). \tag{A.21}$$

Using the Fourier transform of the delta function, we can write

$$\frac{d\delta(x)}{dx} = \delta'(x) = \frac{i}{2\pi}\int_{-\infty}^{+\infty} k e^{ikx} dk. \tag{A.22}$$

Another way of looking at the derivative of the delta function is by means of the following integration by parts of $\delta'(x-a)$:

$$\int_{-\infty}^{\infty} f(x)\delta'(x-a)\,dx = f(x)\delta(x-a)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f'(x)\delta(x-a)\,dx = -f'(a), \tag{A.23}$$

or

$$\int_{-\infty}^{\infty} f(x)\delta'(x-a)\,dx = -f'(a), \tag{A.24}$$

where we have used the fact that $f(x)\delta(x-a)$ is zero at $\pm\infty$. Following the same procedure, we can show that

$$\int_{-\infty}^{\infty} f(x)\delta''(x-a)\,dx = (-1)^2 f''(a) = f''(a). \tag{A.25}$$

Similar repeated integrations by parts lead to the following general relation:

$$\int_{-\infty}^{\infty} f(x)\delta^{(n)}(x-a)\,dx = (-1)^n f^{(n)}(a), \tag{A.26}$$

where $\delta^{(n)}(x-a) = d^n[\delta(x-a)]/dx^n$ and $f^{(n)}(a) = d^n f(x)/dx^n\big|_{x=a}$. In particular, if $f(x) = 1$ and $n = 1$, we have

$$\int_{-\infty}^{\infty} \delta'(x-a)\,dx = 0. \tag{A.27}$$

Here is a list of useful properties of the derivative of the delta function:

$$\delta'(x) = -\delta'(-x), \tag{A.28}$$
$$x\delta'(x) = -\delta(x), \tag{A.29}$$
$$x^2\delta'(x) = 0, \tag{A.30}$$
$$x^2\delta''(x) = 2\delta(x). \tag{A.31}$$

A.2 Three-Dimensional Delta Function

The three-dimensional form of the delta function is given in Cartesian coordinates by

$$\delta(\vec{r}-\vec{r}') = \delta(x-x')\delta(y-y')\delta(z-z') \tag{A.32}$$

and in spherical coordinates by

$$\begin{aligned}\delta(\vec{r}-\vec{r}') &= \frac{1}{r^2}\delta(r-r')\delta(\cos\theta-\cos\theta')\delta(\varphi-\varphi') \\ &= \frac{1}{r^2\sin\theta}\delta(r-r')\delta(\theta-\theta')\delta(\varphi-\varphi'),\end{aligned} \tag{A.33}$$

since, according to (A.17), we have $\delta(\cos\theta - \cos\theta') = \delta(\theta-\theta')/\sin\theta$.

The Fourier transform of the three-dimensional delta function is

$$\delta(\vec{r}-\vec{r}') = \frac{1}{(2\pi)^3}\int d^3k\, e^{i\vec{k}\cdot(\vec{r}-\vec{r}')}, \tag{A.34}$$

and

$$\int d^3r\, f(\vec{r})\delta(\vec{r}) = f(0), \qquad \int d^3r\, f(\vec{r})\delta(\vec{r}-\vec{r}_0) = f(\vec{r}_0). \tag{A.35}$$

The following relations are often encountered:

$$\vec{\nabla}\cdot\left(\frac{\hat{r}}{r^2}\right) = 4\pi\delta(\vec{r}), \qquad \nabla^2\left(\frac{1}{r}\right) = -4\pi\delta(\vec{r}), \tag{A.36}$$

where $\hat{r}$ the unit vector along $\vec{r}$.

We should mention that the physical dimension of the delta function is one over the dimensions of its argument. Thus, if x is a distance, the physical dimension of $\delta(x)$ is given by $[\delta(x)] = 1/[x] = 1/L$, where L is a length. Similarly, the physical dimensions of $\delta(\vec{r})$ is $1/L^3$, since

$$[\delta(x)] = \frac{1}{[x]} = \frac{1}{L} \quad\Longrightarrow\quad [\delta(\vec{r})] = [\delta(x)\delta(y)\delta(z)] = \frac{1}{[x]}\frac{1}{[y]}\frac{1}{[z]} = \frac{1}{L^3}. \tag{A.37}$$

Appendix B

Angular Momentum in Spherical Coordinates

In this appendix, we will show how to derive the expressions of the gradient $\vec{\nabla}$, the Laplacian ∇^2, and the components of the orbital angular momentum in spherical coordinates.

B.1 Derivation of Some General Relations

The Cartesian coordinates (x, y, z) of a vector $\vec{r}$ are related to its spherical polar coordinates (r, θ, ϕ) by

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta. \tag{B.1}$$

The orthonormal Cartesian basis $(\hat{\imath}, \hat{\jmath}, \hat{k})$ is related to its spherical counterpart $(\hat{r}, \hat{\theta}, \hat{\phi})$ by

$$\hat{\imath} = \sin\theta\cos\phi\,\hat{r} + \cos\theta\cos\phi\,\hat{\theta} - \sin\phi\,\hat{\phi}, \tag{B.2}$$

$$\hat{\jmath} = \sin\theta\sin\phi\,\hat{r} + \cos\theta\sin\phi\,\hat{\theta} + \cos\phi\,\hat{\phi}, \tag{B.3}$$

$$\hat{k} = \cos\theta\,\hat{r} - \hat{\theta}\sin\theta. \tag{B.4}$$

Differentiating (B.1), we obtain

$$dx = \sin\theta\cos\phi\,dr + r\cos\theta\cos\phi\,d\theta - r\sin\theta\sin\phi\,d\phi, \tag{B.5}$$

$$dy = \sin\theta\sin\phi\,dr + r\cos\theta\sin\phi\,d\theta + r\sin\theta\cos\phi\,d\phi, \tag{B.6}$$

$$dz = \cos\theta\,dr - r\sin\theta\,d\theta. \tag{B.7}$$

Solving these equations for dr, $d\theta$, and $d\phi$, we obtain

$$dr = \sin\theta\cos\phi\,dx + \sin\theta\sin\phi\,dy + \cos\theta\,dz, \tag{B.8}$$

$$d\theta = \frac{1}{r}\cos\theta\cos\phi\,dx + \frac{1}{r}\cos\theta\sin\phi\,dy - \frac{1}{r}\sin\theta\,dz, \tag{B.9}$$

$$d\phi = -\frac{\sin\phi}{r\sin\theta}\,dx + \frac{\cos\phi}{r\sin\theta}\,dy. \tag{B.10}$$

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

We can verify that (B.5) to (B.10) lead to

$$\frac{\partial r}{\partial x} = \sin\theta\cos\phi, \qquad \frac{\partial \theta}{\partial x} = \frac{1}{r}\cos\phi\cos\theta, \qquad \frac{\partial \phi}{\partial x} = -\frac{\sin\phi}{r\sin\theta}, \tag{B.11}$$

$$\frac{\partial r}{\partial y} = \sin\theta\sin\phi, \qquad \frac{\partial \theta}{\partial y} = \frac{1}{r}\sin\phi\cos\theta, \qquad \frac{\partial \phi}{\partial y} = \frac{\cos\phi}{r\sin\theta}, \tag{B.12}$$

$$\frac{\partial r}{\partial z} = \cos\theta, \qquad \frac{\partial \theta}{\partial z} = -\frac{1}{r}\sin\theta, \qquad \frac{\partial \phi}{\partial z} = 0, \tag{B.13}$$

which, in turn, yield

$$\begin{aligned}\frac{\partial}{\partial x} &= \frac{\partial}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial}{\partial \theta}\frac{\partial \theta}{\partial x} + \frac{\partial}{\partial \phi}\frac{\partial \phi}{\partial x} \\ &= \sin\theta\cos\phi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\phi\frac{\partial}{\partial \theta} - \frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial \phi},\end{aligned} \tag{B.14}$$

$$\begin{aligned}\frac{\partial}{\partial y} &= \frac{\partial}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial}{\partial \theta}\frac{\partial \theta}{\partial y} + \frac{\partial}{\partial \phi}\frac{\partial \phi}{\partial y} \\ &= \sin\theta\sin\phi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\phi\frac{\partial}{\partial \theta} + \frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial \phi},\end{aligned} \tag{B.15}$$

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial r}\frac{\partial r}{\partial z} + \frac{\partial}{\partial \theta}\frac{\partial \theta}{\partial z} + \frac{\partial}{\partial \phi}\frac{\partial \phi}{\partial z} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}. \tag{B.16}$$

B.2 Gradient and Laplacian in Spherical Coordinates

We can show that a combination of (B.14) to (B.16) allows us to express the operator $\vec{\nabla}$ in spherical coordinates:

$$\vec{\nabla} = \hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z} = \hat{r}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial \theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}, \tag{B.17}$$

and also the Laplacian operator ∇^2:

$$\nabla^2 = \vec{\nabla}.\vec{\nabla} = \left(\hat{r}\frac{\partial}{\partial r} + \frac{\hat{\theta}}{r}\frac{\partial}{\partial \theta} + \frac{\hat{\phi}}{r\sin\phi}\frac{\partial}{\partial \phi}\right)\cdot\left(\hat{r}\frac{\partial}{\partial r} + \frac{\hat{\theta}}{r}\frac{\partial}{\partial \theta} + \frac{\hat{\phi}}{r\sin\theta}\frac{\partial}{\partial \phi}\right). \tag{B.18}$$

Now, using the relations

$$\frac{\partial \hat{r}}{\partial r} = 0, \qquad \frac{\partial \hat{\theta}}{\partial r} = 0, \qquad \frac{\partial \hat{\phi}}{\partial r} = 0, \tag{B.19}$$

$$\frac{\partial \hat{r}}{\partial \theta} = \hat{\theta}, \qquad \frac{\partial \hat{\theta}}{\partial \theta} = -\hat{r}, \qquad \frac{\partial \hat{\phi}}{\partial \theta} = 0, \tag{B.20}$$

$$\frac{\partial \hat{r}}{\partial \phi} = \hat{\phi}\sin\theta, \qquad \frac{\partial \hat{\theta}}{\partial \phi} = \hat{\phi}\cos\theta, \qquad \frac{\partial \hat{\phi}}{\partial \phi} = -\hat{r}\sin\theta - \hat{\theta}\cos\theta, \tag{B.21}$$

we can show that the Laplacian operator reduces to

$$\nabla^2 = \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial}{\partial \theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial \phi^2}\right]. \tag{B.22}$$

B.3 Angular Momentum in Spherical Coordinates

The orbital angular momentum operator $\vec{L}$ can be expressed in spherical coordinates as

$$\hat{\vec{L}} = \hat{\vec{R}} \times \hat{\vec{P}} = (-i\hbar r)\hat{r} \times \vec{\nabla} = (-i\hbar r)\hat{r} \times \left[\hat{r}\frac{\partial}{\partial r} + \frac{\hat{\theta}}{r}\frac{\partial}{\partial \theta} + \frac{\hat{\phi}}{r\sin\theta}\frac{\partial}{\partial \phi}\right], \tag{B.23}$$

or as

$$\hat{\vec{L}} = -i\hbar\left(\hat{\phi}\frac{\partial}{\partial \theta} - \frac{\hat{\theta}}{\sin\theta}\frac{\partial}{\partial \phi}\right). \tag{B.24}$$

Using (B.24) along with (B.2) to (B.4), we express the components $\hat{L}_x, \hat{L}_y, \hat{L}_z$ within the context of the spherical coordinates. For instance, the expression for $\hat{L}_x$ can be written as follows:

$$\begin{aligned}\hat{L}_x &= \hat{\imath}.\hat{\vec{L}} = -i\hbar\left(\hat{r}\sin\theta\cos\phi + \hat{\theta}\cos\theta\cos\phi - \hat{\phi}\sin\phi\right)\cdot\left(\hat{\phi}\frac{\partial}{\partial\theta} - \frac{\hat{\theta}}{\sin\theta}\frac{\partial}{\partial\phi}\right)\\ &= i\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right).\end{aligned} \tag{B.25}$$

Similarly, we can easily obtain

$$\hat{L}_y = i\hbar\left(-\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right), \tag{B.26}$$

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}. \tag{B.27}$$

From the expressions (B.25) and (B.26) for $\hat{L}_x$ and $\hat{L}_y$, we infer that

$$\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y = \pm\hbar e^{\pm i\phi}\left(\frac{\partial}{\partial\theta} \pm i\cot\theta\frac{\partial}{\partial\phi}\right). \tag{B.28}$$

The expression for $\vec{L}^2$ is

$$\vec{L}^2 = -\hbar^2 r^2(\hat{r}\times\vec{\nabla})\cdot(\hat{r}\times\vec{\nabla}) = -\hbar^2 r^2\left[\nabla^2 - \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)\right]; \tag{B.29}$$

it can be easily written in terms of the spherical coordinates as

$$\vec{L}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]. \tag{B.30}$$

This expression was derived by substituting (B.22) into (B.29).

Note that, using the expression (B.29) for $\vec{L}^2$, we can rewrite ∇^2 as

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{\hbar^2 r^2}\vec{L}^2 = \frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{1}{\hbar^2 r^2}\vec{L}^2. \tag{B.31}$$

Appendix C

C++ Code for Solving the Schrödinger Equation

C.1 Introduction

In this appendix, we deal with the numerical solution of the Schrödinger equation. For this, we are going to develop a C++ code which is designed to solve the one-dimensional Schrödinger equation for a harmonic oscillator (HO) potential as well as for an infinite square well (ISW) potential as outlined in Chapter 4. My special thanks are due to Dr. M. Bulut and to Prof. Dr. H. Mueller-Krumbhaar and his Ph.D. student C. Gugenberger who have worked selflessly hard to write and test the code listed below. Dr. Mevlut wrote an early code for the ISW, while Prof. Mueller-Krumbhaar and Gugenberger not only wrote a new code (see the version listed below) for the HO but also designed it in a way that it applies to the ISW potential as well (they have also added effective didactic comments so that our readers can effortlessly understand the code and make use of it).

Note: to shift from the harmonic oscillator code to the infinite square well code, one needs simply to erase the first double forward-slash (i.e., "//") from the oscillator's program line below:

```
E_pot[i] = 0.5*dist*dist; // E_pot[i]=0;//E_pot=0:Infinite Well!
```

Of course, one still needs to rescale the energy and the value of 'xRange' in order to agree with the algorithm outlined at the end of Chapter 4.

C.2 C++ Code for Solving the Schrödinger Equation

The C++ Code: osci.cpp

```
/* osci.cpp: Solution of the one-dimensional Schrodinger equation for
  a particle in a harmonic potential, using the shooting method.
  To compile and link with gnu compiler, type: g++ -o osci osci.cpp
  To run the current C++ program, simply type: osci
```

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

```
  Plot by gnuplot: /GNUPLOT> set terminal windows
                   /GNUPLOT> plot "psi-osc.dat" with lines  */
#include <cstdio>
#include <cstdlib>
#include <cmath>
#define MAX(a, b) (((a) > (b)) ? (a) : (b))
int main(int argc, char*argv[])
{// Runtime constants
const static double Epsilon = 1e-10; // Defines the precision of
                //...  energy calculations
const static int N_of_Divisions = 1000;
const static int N_max = 5; //Number of calculated Eigenstates

FILE *Wavefunction_file, *Energy_file, *Potential_file;
Wavefunction_file = fopen("psi-osc.dat","w");
Energy_file = fopen("E_n_Oszillator.dat","w");
Potential_file = fopen("HarmonicPotentialNoDim.dat", "w");
if (!(Wavefunction_file && Energy_file && Potential_file))
{ printf("Problems to create files output.\n"); exit(2); }

/* Physical parameters using dimensionless quantities.
 ATTENTION:  We set initially:  hbar = m = omega = a = 1, and
 reintroduce physical values at the end.  According to Eq.(4.117),
 the ground state energy then is  E_n = 0.5. Since the wave function
 vanishes only at -infinity and  +infinity, we have to cut off the
 calculation somewhere, as given by  'xRange'. If xRange is chosen
 too large, the open (positive) end of  the  wave function can
 diverge numerically in this simple shooting  approach. */

const static double xRange = 12; // xRange=11.834 corresponds to a
  //... physical range of -20fm < x < +20fm, see after Eq.(4.199).
const static double h_0 = xRange / N_of_Divisions;
double* E_pot = new double[N_of_Divisions+1];
double dist;

for (int i = 0; i <= N_of_Divisions; ++i)
{ // Harmonic potential, as given in Eq. (4.115), but dimensionless
dist = i*h_0 - 0.5*xRange;
E_pot[i] = 0.5*dist*dist; // E_pot[i]=0;//E_pot=0:Infinite Well!
fprintf(Potential_file, "%16.12e \t\t %16.12e\n", dist, E_pot[i]);
}
fclose(Potential_file);

/* Since the Schrodinger equation is linear, the amplitude of the
 wavefunction will be fixed by normalization.
At left we set it small but nonzero. */
const static double Psi_left = 1.0e-3; // left boundary condition
```

```
const static double Psi_right = 0.0; // right boundary condition

double *Psi, *EigenEnergies;// Arrays to hold the results
Psi = new double[N_of_Divisions+1]; //N_of_Points = N_of_Divisions+1
EigenEnergies = new double[N_max+1];
Psi[0] = Psi_left;
Psi[1] = Psi_left + 1.0e-3; // Add arbitrary small value

int N_quantum;//N_quantum is Energy Quantum Number
int Nodes_plus; // Number of nodes (+1) in wavefunction
double K_square;// Square of wave vector
// Initial Eigen-energy search limits
double E_lowerLimit = 0.0;// Eigen-energy must be positive
double E_upperLimit = 10.0;
int End_sign = -1;
bool Limits_are_defined = false;
double Normalization_coefficient;
double E_trial;

// MAIN LOOP begins:-----------------------------------
for(N_quantum=1; N_quantum <= N_max; ++N_quantum)
{
// Find the eigen-values for energy. See theorems (4.1) and (4.2).
Limits_are_defined = false;
while (Limits_are_defined == false)
{ /* First, determine an upper limit for energy, so that the wave-
 function Psi[i] has one node more than physically needed.    */
Nodes_plus = 0;
E_trial = E_upperLimit;
for (int i=2; i <= N_of_Divisions; ++i)
{ K_square = 2.0*(E_trial - E_pot[i]);
// Now use the NUMEROV-equation (4.197) to calculate wavefunction
Psi[i] = 2.0*Psi[i-1]*(1.0 - (5.0*h_0*h_0*K_square / 12.0))
/(1.0 + (h_0*h_0*K_square/12.0))-Psi[i-2];
if (Psi[i]*Psi[i-1] < 0) ++Nodes_plus;
}
/* If one runs into the following condition, the modification
 of the upper limit was too aggressive. */
if (E_upperLimit < E_lowerLimit)
E_upperLimit = MAX(2*E_upperLimit, -2*E_upperLimit);
if (Nodes_plus > N_quantum) E_upperLimit *= 0.7;
else if (Nodes_plus < N_quantum) E_upperLimit *= 2.0;
else Limits_are_defined = true; // At least one node should appear.
} // End of the loop: while (Limits_are_defined == false)
// Refine the energy by satisfying the right boundary condition.
End_sign = -End_sign;
while ((E_upperLimit - E_lowerLimit) > Epsilon)
```

```
{ E_trial = (E_upperLimit + E_lowerLimit) / 2.0;
for (int i=2; i <= N_of_Divisions; ++i)
{ // Again eq.(4.197) of the Numerov-algorithm:
K_square = 2.0*(E_trial - E_pot[i]);
Psi[i] = 2.0*Psi[i-1] * (1.0 - (5.0*h_0*h_0*K_square / 12.0))
 /(1.0 + (h_0*h_0*K_square/12.0))-Psi[i-2];
}
if (End_sign*Psi[N_of_Divisions] > Psi_right) E_lowerLimit = E_trial;
else E_upperLimit = E_trial;
} // End of loop: while ((E_upperLimit - E_lowerLimit) > Epsilon)

// Initialization for the next iteration in main loop
E_trial = (E_upperLimit+E_lowerLimit)/2;
EigenEnergies[N_quantum] = E_trial;
E_upperLimit = E_trial;
E_lowerLimit = E_trial;

// Now find the normalization coefficient
double Integral = 0.0;
for (int i=1; i <= N_of_Divisions; ++i)
{ // Simple integration
Integral += 0.5*h_0*(Psi[i-1]*Psi[i-1]+Psi[i]*Psi[i]);
}
Normalization_coefficient = sqrt(1.0/Integral);
// Output of normalized dimensionless wave function
for (int i=0; i <=N_of_Divisions; ++i)
{ fprintf(Wavefunction_file, "%16.12e \t\t %16.12e\n",
 i*h_0 - 0.5*xRange, Normalization_coefficient*Psi[i]);
}
fprintf(Wavefunction_file,"\n");
} // End of MAIN LOOP. --------------------------------
fclose(Wavefunction_file);

/*Finally convert dimensionless units in real units. Note that
  energy does not depend explicitly on the particle's mass anymore:
    hbar = 1.05457e-34;// Planck constant/2pi
    omega = 5.34e21; // Frequency in 1/s
    MeV = 1.602176487e-13; // in J
    The correct normalization would be hbar*omega/MeV = 3.5148461144,
    but we use the approximation 3.5 for energy-scale as in chap. 4.9 */
const static double Energyscale = 3.5;// in MeV
// Output with rescaled dimensions; assign Energy_file
printf("Quantum Harmonic Oscillator, program osci.cpp\n");
printf("Energies in MeV:\n");
printf("n \t\t E_n\n");
for (N_quantum=1; N_quantum <= N_max; ++N_quantum)
{ fprintf(Energy_file,"%d \t\t %16.12e\n", N_quantum-1,
```

```
  Energyscale*EigenEnergies[N_quantum]);
  printf("%d \t\t %16.12e\n", N_quantum-1,
  Energyscale*EigenEnergies[N_quantum]);
}
fprintf(Energy_file,"\n");
fclose(Energy_file);
printf("Wave-Functions in File: psi_osc.dat \n");
printf("\n");
return 0;
}
```

Appendix D

Index Notation for 4-Vectors

Index notation is a powerful tool that greatly simplifies the algebra involved in dealing with vectors and matrices; it is far superior than the standard vector notation. As we will see below, the index notation will allows us to derive or establish many vector identities as well as carry out algebraic manipulations in a far more transparent and simpler manner than the conventional vector notation.

Before presenting the index notation for 4-dimensional vectors, let us first take a look at the index notation of the familiar case of 3-dimensional vectors in the Euclidean space.

D.1 Index Notation for 3 and 4-Vectors

D.1.1 Introductory Reminder

Consider a point P in the 3-dimensional Euclidean space. The position vector of Point P in the Cartesian, cylindrical, and spherical coordinate systems is given, respectively, by

$$\text{Cartesian:} \qquad \vec{r} = x\hat{\imath} + y\hat{\jmath} + z\hat{k}, \tag{D.1}$$

$$\text{Cylindrical:} \qquad \vec{r} = \rho\cos\phi\,\hat{\imath} + \rho\sin\phi\,\hat{\jmath} + z\hat{k}, \tag{D.2}$$

$$\text{Spherical:} \qquad \vec{r} = r\sin\theta\cos\phi\,\hat{\imath} + r\sin\theta\sin\phi\,\hat{\jmath} + r\cos\theta\,\hat{k}, \tag{D.3}$$

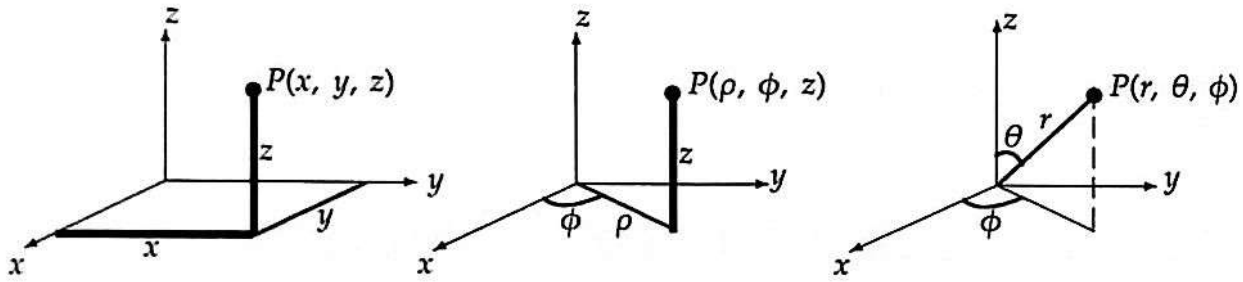
where, as illustrated in Figure D.1, the cylindrical, (ρ, ϕ, z), and spherical, (r, θ, ϕ), coordinates are related to their Cartesian, (x, y, z), as follows:

$$\rho = \sqrt{x^2 + y^2}, \quad r = \sqrt{x^2 + y^2 + z^2}, \quad \phi = \tan^{-1}\left(\frac{y}{x}\right), \quad \theta = \tan^{-1}\left(\frac{\sqrt{x^2 + y^2}}{z}\right), \quad z = z. \tag{D.4}$$

The basis vectors in the Cartesian, cylindrical, and spherical coordinate systems are given, respectively, by

$$\left(\hat{\imath}, \hat{\jmath}, \hat{k}\right), \qquad \left(\hat{\rho}, \hat{\phi}, \hat{z}\right), \qquad \left(\hat{r}, \hat{\theta}, \hat{\phi}\right). \tag{D.5}$$

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.



(a) Cartesian coordinates (b) Cylindrical coordinates (c) Spherical coordinates

Figure D.1: Cartesian, Cylindrical, and Spherical coordinates

Each of these sets of unit vectors form an orthonormal basis

$$\hat{\imath}\cdot\hat{\jmath}=\hat{\imath}\cdot\hat{k}=\hat{\jmath}\cdot\hat{k}=0, \qquad \hat{\imath}^2=\hat{\jmath}^2=\hat{k}^2=1, \tag{D.6}$$

$$\hat{\rho}\cdot\hat{\phi}=\hat{\rho}\cdot\hat{z}=\hat{\phi}\cdot\hat{z}=0, \qquad \hat{\rho}^2=\hat{\phi}^2=\hat{z}^2=1, \tag{D.7}$$

$$\hat{r}\cdot\hat{\theta}=\hat{r}\cdot\hat{\phi}=\hat{\theta}\cdot\hat{\phi}=0, \qquad \hat{r}^2=\hat{\theta}^2=\hat{\phi}^2=1. \tag{D.8}$$

We can express the cylindrical basis vectors in terms of their Cartesian counterparts as follows:

$$\hat{\rho}=\cos\phi\,\hat{\imath}+\sin\phi\,\hat{\jmath}, \qquad \hat{\phi}=-\sin\phi\,\hat{\imath}+\cos\phi\,\hat{\jmath}, \qquad \hat{k}=\hat{k}. \tag{D.9}$$

We can also express the spherical basis vectors in terms of their Cartesian counterparts as follows:

$$\hat{r} = \sin\theta\cos\phi\,\hat{\imath}+\sin\theta\sin\phi\,\hat{\jmath}+\cos\theta\,\hat{k}, \tag{D.10}$$

$$\hat{\theta} = \cos\theta\cos\phi\,\hat{\imath}+\cos\theta\sin\phi\,\hat{\jmath}-\sin\theta\,\hat{k}, \tag{D.11}$$

$$\hat{\phi} = -\sin\phi\,\hat{\imath}+\cos\phi\,\hat{\jmath}. \tag{D.12}$$

It is worth mentioning that the position vector $\vec{r}$ (or any other vector) exists independently of the coordinate system we use to describe it. For instance, we can describe a vector $\vec{A}$ by means of its Cartesian, (A_x, A_y, A_z), cylindrical, (A_ρ, A_ϕ, A_z), or spherical, (A_r, A_θ, A_ϕ), components as follows:

$$\text{Cartesian:} \qquad \vec{A}=A_x\,\hat{\imath}+A_y\,\hat{\jmath}+A_z\,\hat{k}, \tag{D.13}$$

$$\text{Cylindrical:} \qquad \vec{A}=A_\rho\,\hat{\rho}+A_\phi\,\hat{\phi}+A_z\,\hat{k}, \tag{D.14}$$

$$\text{Spherical:} \qquad \vec{A}=A_r\,\hat{r}+A_\theta\,\hat{\theta}+A_\phi\,\hat{\phi}. \tag{D.15}$$

We can express the scalar (inner) product, $\vec{A}\cdot\vec{B}$, between two arbitrary vectors $\vec{A}$ and $\vec{B}$ in the Cartesian, cylindrical, and spherical coordinate systems as follows

$$\text{Cartesian:} \qquad \vec{A}\cdot\vec{B}=A_xB_x+A_yB_y+A_zB_z, \tag{D.16}$$

$$\text{Cylindrical:} \qquad \vec{A}\cdot\vec{B}=A_\rho B_\rho+A_\phi B_\phi+A_zB_z, \tag{D.17}$$

$$\text{Spherical:} \qquad \vec{A}\cdot\vec{B}=A_rB_r+A_\theta B_\theta+A_\phi B_\phi. \tag{D.18}$$

D.1.2 Index Notation for 3-Vectors

At this level, we want to introduce the **index notation** which is known to greatly simplify vector calculus. This notation will free us from committing to any particular representation. That is, instead of restrictively using any of the Cartesian, cylindrical, or spherical bases – $(\hat{i}, \hat{j}, \hat{k})$, $(\hat{\rho}, \hat{\phi}, \hat{z})$, $(\hat{r}, \hat{\theta}, \hat{\phi})$ – we are going to introduce a more convenient and more general orthonormal basis $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$ that spans any 3-dimensional Euclidean space:

$$(\hat{i}, \hat{j}, \hat{k}), \quad (\hat{\rho}, \hat{\phi}, \hat{z}), \quad (\hat{r}, \hat{\theta}, \hat{\phi}) \quad \longrightarrow \quad (\hat{e}_1, \hat{e}_2, \hat{e}_3). \tag{D.19}$$

Hence, the index notation has the advantage of freeing us from restricting ourselves to any particular coordinate system; the basis vectors $\hat{e}_i$ (where $i = 1, 2, 3$) are general and apply to any 3-dimensional coordinate system (Cartesian, cylindrical, spherical, etc.). For instance, instead of writing an arbitrary vector $\vec{A}$ in terms of its Cartesian, cylindrical, or spherical components (as shown in equations (D.13)–(D.15)), we will write it in the following more general and simple form that is valid in any 3-dimensional coordinate system:

$$(A_x, A_y, A_z), \quad (A_\rho, A_\phi, A_z), \quad (A_r, A_\theta, A_\phi) \quad \longrightarrow \quad (A_1, A_2, A_3); \tag{D.20}$$

that is, we can write the vector $\vec{A}$ in the index notation as follows

$$\vec{A} = A_1\hat{e}_1 + A_2\hat{e}_2 + A_3\hat{e}_3 = \sum_{i=1}^{3} A_i\hat{e}_i \equiv A_i\hat{e}_i. \tag{D.21}$$

In the last step of this relation, we have used the *Einstein summation convention* which does away with the explicit appearance of the summation symbol. According to this convention, whenever an index variable (here "i") appears twice in a single term, summation of that term over all the values of the index is automatically implied (here $i = 1, 2, 3$), even though the summation symbol is not explicitly written. This is one of the advantages for using the index notation. For notational brevity, we will be using the Einstein summation convention throughout Chapter 12 and the relevant appendices, D, E, F, and G.

In the case of a Cartesian coordinate system, Equation (D.21) implies that index 1 stands for x, 2 for y, and 3 for z; that is, $A_1 = x$, $A_2 = y$, $A_3 = z$ and $\hat{e}_1 = \hat{\imath}$, $\hat{e}_1 = \hat{\jmath}$, $\hat{e}_1 = \hat{k}$. In this way, the ith component of a vector $\vec{A}$ is denoted by A_i.

Instead of the cumbersome relations (D.6)–(D.8), we can express the orthonormality condition of the basis $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$ in the following simpler form

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}, \qquad i, j = 1, 2, 3., \tag{D.22}$$

where δ_{ij} is the 3-dimensional *Kronecker delta*. Note that, using the Einstein summation convention, we can write

$$\delta_{ii} = 3, \qquad \delta_{ij}\delta_{jk} = \delta_{ik}, \qquad \delta_{ij}\delta_{jk}\delta_{ki} = \delta_{ik}\delta_{ki} = \delta_{ii} = 3, \qquad i, j, k = 1, 2, 3. \tag{D.23}$$

Another notable usefulness of the index notation is the fact that it lends itself quite readily to the matrix representation, where basis vectors (or any vector $\vec{A}$) are represented by column

matrices and the Kronecker delta by a 3×3 matrix[1]

$$(\hat{e}_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad (\hat{e}_2) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad (\hat{e}_3) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad (\vec{A}) = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}, \quad [\delta] = (\delta_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{D.24}$$

Clearly, $[\delta]$ is nothing but the 3-dimensional identity matrix; we will denote it by I_3. We should note that, in vector calculus, δ_{ij} is also known as the lower version of the Euclidean metric tensor g_{ij} where the **Euclidean metric** $[g]$ is given in matrix form[2] by

$$[g] = (g_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv [\delta] = (\delta_{ij}). \tag{D.25}$$

Note: In this text, we use $[g]$ or (g_{ij}) to denote the matrix of the metric, while g_{ij} represents the ij element of the metric; the same notation for the Kronecker delta function where $[\delta]$ or (δ_{ij}) represent the matrix and δ_{ij} the ij element of the $\delta-$matrix.

Any component A_i of the vector $\vec{A}$ is given by

$$A_i = \hat{e}_i \cdot \vec{A}, \qquad i = 1, 2, 3. \tag{D.26}$$

We should note that the property $g_{ij} = \delta_{ij}$ holds here only because we are working in a flat-space coordinate system (no curvature); this is particularly the case here because the Euclidean space is flat. We are going to discuss the impact of the space's curvature on the metric in Section D.2

D.1.3 Inner and Cross Products of 3-Vectors in the Index Notation

Within the index notation, we can easily perform the inner product between two vectors $\vec{A} = \sum_{i=1}^{3} A_i \hat{e}_i$ and $\vec{B} = \sum_{i=1}^{3} B_i \hat{e}_i$:

$$\begin{aligned} \vec{A} \cdot \vec{B} &= \left(\sum_{i=1}^{3} A_i \hat{e}_i \right) \cdot \left(\sum_{j=1}^{3} B_j \hat{e}_j \right) = \sum_{i=1}^{3} \sum_{j=1}^{3} A_i B_j \left(\hat{e}_i \cdot \hat{e}_j \right) = \sum_{i=1}^{3} \sum_{j=1}^{3} A_i B_j g_{ij} \\ &\equiv \sum_{i=1}^{3} A_i B_i = A_1 B_1 + A_2 B_2 + A_3 B_3. \end{aligned} \tag{D.27}$$

That is, we can write the inner product between two vectors $\vec{A} = A_i \hat{e}_i$ and $\vec{B} = B_i \hat{e}_i$ in the index notation in the following more compact form

$$\boxed{\vec{A} \cdot \vec{B} = A_i B_i, \qquad i = 1, 2, 3.} \tag{D.28}$$

[1] We will be using $(\hat{e}_i)$ to denote the column matrices corresponding to the unit vectors $\hat{e}_i$ and δ or (δ_{ij}) to denote the 3×3 matrix corresponding to the Kronecker δ as displayed in (D.24); moreover, we will use the notation δ_{ij} to correspond to the ij matrix element of δ.

[2] In this text, we use $[g]$ or (g_{ij}) to denote the matrix of the metric, while g_{ij} represents the ij element of the metric; the same notation for the Kronecker delta function where $[\delta]$ or (δ_{ij}) represent the matrix and δ_{ij} the ij element of the $\delta-$matrix.

If the coordinate system, which is defined by the three basis vectors $\hat{e}_i$ (where $i = 1, 2, 3$), is rotated by a certain angle θ, the new system is defined by the new three basis vectors $\hat{e}'_1$, $\hat{e}'_2$, $\hat{e}'_3$ which obey the same relation as (D.22), namely

$$\hat{e}'_i \cdot \hat{e}'_j = \delta_{ij}, \qquad i, j = 1,2,3. \tag{D.29}$$

Using (D.22) and (D.29), we can express the inner product between the basis vectors in terms of the Euclidean metric g as follows:

$$\hat{e}_i \cdot \hat{e}_j = g_{ij}, \qquad \hat{e}'_i \cdot \hat{e}'_j = g_{ij}, \qquad i, j = 1,2,3. \tag{D.30}$$

Clearly, the Euclidean metric g is symmetric, $g_{ij} = g_{ji}$, since the inner product between the basis vectors does not depend on their order in the product.

The vectors $\vec{A}$ and $\vec{B}$ are given in the *rotated* coordinate system by

$$\vec{A}' = \sum_{i=1}^{3} A'_i \hat{e}'_i = A'_i \hat{e}'_i, \qquad \vec{B}' = \sum_{i=1}^{3} B'_j \hat{e}'_j = B'_j \hat{e}'_j, \tag{D.31}$$

and their inner product by

$$\vec{A}' \cdot \vec{B}' = \left(A'_i \hat{e}'_i\right) \cdot \left(B'_j \hat{e}'_j\right) = A'_i B'_j \left(\hat{e}'_i \cdot \hat{e}'_j\right) = A'_i B'_j g_{ij} = A'_i B'_i, \tag{D.32}$$

where use was made of (D.30) and where g_{ij} is given by (D.25).

As shown in Chapter 7 (see Section 7.1), the inner product is invariant[3] under rotations; namely,

$$\boxed{A_1 B_1 + A_2 B_2 + A_3 B_3 = A'_1 B'_1 + A'_2 B'_2 + A'_3 B'_3 \qquad \text{or} \qquad A_i B_i = A'_i B'_i.} \tag{D.33}$$

This implies that the inner product between two vectors is an invariant under rotations; in particular, the inner product $\vec{r} \cdot \vec{r}$ is also an invariant:

$$\vec{r}^{\,2} = \vec{r}'^{\,2} \qquad \longrightarrow \qquad x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2. \tag{D.34}$$

The cross product (between two vectors $\vec{A}$ and $\vec{B}$) is given by

$$\vec{A} \times \vec{B} = (A_i \hat{e}_i) \times (B_j \hat{e}_j) = (A_2 B_3 - A_3 B_2)\,\hat{e}_1 + (A_3 B_1 - A_1 B_3)\,\hat{e}_2 + (A_1 B_2 - A_2 B_1)\,\hat{e}_3, \tag{D.35}$$

and can be written in a much concise and simpler form in the index notation as follows

$$\boxed{\vec{A} \times \vec{B} = A_i B_j \left(\hat{e}_i \times \hat{e}_j\right) \equiv \epsilon_{ijk} A_i B_j \hat{e}_k, \qquad i, j, k = 1, 2, 3,} \tag{D.36}$$

where we have used the following relation

$$\boxed{\hat{e}_i \times \hat{e}_j = \epsilon_{ijk} \hat{e}_k, \qquad i, j, k = 1, 2, 3,} \tag{D.37}$$

[3] A quantity is said to be invariant if it has the same numerical values in all inertial frames (frames that have no relative acceleration). Similarly, a relation (or equation) is invariant if it has the same *form* in all inertial frames.

where ϵ_{ijk} is the *Levi-Civita tensor* which can take only three values -1, 0, 1 according to the following rule:

$$\epsilon_{ijk} = \begin{cases} -1, & \text{if } i\,j\,k \text{ are odd permutations of } 1\,2\,3 \quad (i.e., \epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1), \\ 1, & \text{if } i\,j\,k \text{ are even permutations of } 1\,2\,3 \quad (i.e., \epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1), \\ 0, & \text{If any two or more indices are equal} \quad (i.e., \epsilon_{111} = \epsilon_{222} = \epsilon_{333} = \epsilon_{112} = \cdots = 0). \end{cases} \tag{D.38}$$

It is worth mentioning that the following practical relation also yields the correct values for ϵ_{ijk}:

$$\epsilon_{ijk} = \frac{1}{2}(i-j)(j-k)(k-i), \qquad i,\ j,\ k = 1,\ 2,\ 3. \tag{D.39}$$

The Levi-Civita tensor is usually written in terms of the product of Kronecker deltas as follows:

$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{kl}\delta_{jm}, \qquad i,\ j,\ k,\ l,\ m = 1,\ 2,\ 3. \tag{D.40}$$

We can also write

$$\epsilon_{ijk}\epsilon_{ijl} = 2\delta_{kl} \qquad \epsilon_{ijk}\epsilon_{ijk} = 2\delta_{ii} = 6 \qquad i,\ j,\ k,\ l = 1,\ 2,\ 3. \tag{D.41}$$

Vector Calculus and the Index Notation

The index notation easily allows for expressing quantities or operators such as the divergence or the curl of a vector in a very simple manner. Using the notation $\partial_i = \partial/\partial x_i$, we can write the divergence of a vector $\vec{A}$ as follows:

$$\begin{aligned} \vec{\nabla}\cdot\vec{A} &= \frac{\partial A_1}{\partial x_1} + \frac{\partial A_2}{\partial x_2} + \frac{\partial A_3}{\partial x_3} \equiv \frac{\partial A_i}{\partial x_i} \\ &= \partial_i A_i, \end{aligned} \tag{D.42}$$

where $x_1 = x, x_2 = y, x_3 = z$. We can also express the curl of a vector $\vec{A}$ in the following concise form:

$$\begin{aligned} \vec{\nabla}\times\vec{A} &= (\partial_2 A_3 - \partial_3 A_2)\hat{e}_1 + (\partial_3 A_1 - \partial_1 A_3)\hat{e}_2 + (\partial_1 A_2 - \partial_2 A_1)\hat{e}_3 \\ &\equiv \epsilon_{ijk}\partial_i A_j \hat{e}_k, \qquad i,\ j,\ k = 1,\ 2,\ 3. \end{aligned} \tag{D.43}$$

We can also express the Laplacian operator as follows:

$$\Delta \equiv \nabla^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}. \tag{D.44}$$

D.1.4 Index Notation for 4-Vectors

Unlike Newtonian mechanics which functions well within the 3-dimensional Euclidean space, the theory of special relativity requires working in a 4-dimensional space. The Minkowski spacetime provides the best mathematical framework to describe events in special relativity where physical events are described by four components: one time component and three spatial components where the time and space components are treated on an equal footing. For instance, the position of an event in the Minkowski spacetime is described by a 4-vector: $x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \vec{r})$ where all of the four components have the same

physical unit of length. Four-vector are written with Greek superscripts (here μ); for instance, we can write any general 4-vector A^μ as follows

$$A^\mu = (A^0, A^1, A^2, A^3) \equiv (A^0, \vec{A}) \equiv (A^0, A^i), \qquad \mu = 0,1,2,3 \quad \text{and} \quad i = 1,2,3. \tag{D.45}$$

To describe events in spacetime, we need to introduce a basis of four unit vectors $(\hat{e}_0, \hat{e}_1, \hat{e}_2, \hat{e}_3)$ that point along the t, x, y, z axes. By analogy to (D.21), any 4-vector A^μ may be expressed in the following form

$$A^\mu = A^0\hat{e}_0 + A^1\hat{e}_1 + A^2\hat{e}_2 + A^3\hat{e}_3 = \sum_{\nu=0}^{3} A^\nu \hat{e}_\nu \equiv A^\nu \hat{e}_\nu, \tag{D.46}$$

where the Einstein summation convention is used in the last step of the previous equation. We will use Greek letters (μ, ν, etc.,) to denote spacetime indices and Latin letters (i, j, k, etc.,) for spatial indices; Greek indices run over the values $0, 1, 2, 3$, while Latin indices over the values $1, 2, 3$.

D.2 The Minkowski Metric

The metric is the relation that gives the distance between two points in a particular space: once the space's metric is determined, we know practically everything about the geometry of that space. The geometry of the space, and hence the metric, depends on the strength of gravity; the metric is affected by the absence or presence of gravity. According to the theory of General Relativity, gravitation is viewed as a *curved* spacetime phenomenon where one has to consider the deformations induced by matter (gravity) on the space's geometry. However, in the case of weak (or no) gravity, spacetime can be viewed as *flat*; flat spacetime, often represented by the 4-dimensional Minkowski spacetime, can be described best by Special Relativity. So, *the Minkowski spacetime is used to describe physical systems over finite distances where gravity is not significant*. This is the underlying reason for often viewing the Minkowski space as *flat spacetime*. The Minkowskian description is just an approximation which is valid on (relatively) local portions of spacetime. In sum, when gravity is significant, spacetime becomes *curved*; in this case, one would need to abandon Special Relativity and, instead, use General Relativity.

Consider two infinitesimally separated spacetime events that are located at

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z), \quad \text{and} \quad x^\mu + dx^\mu = (ct + cdt, x + dx, y + dy, z + dz). \tag{D.47}$$

In the 4-dimensional spacetime, where the basis vectors are represented by $(\hat{e}_0, \hat{e}_1, \hat{e}_2, \hat{e}_3)$, the infinitesimal displacement vector is given by

$$ds = \sum_{\mu=0}^{3} \sqrt{f_{x^\mu}} dx^\mu \hat{e}_\mu \equiv \sqrt{f_{x^\mu}} dx^\mu \hat{e}_\mu, \tag{D.48}$$

where f_{x^μ} are scale factors that depend on the coordinates x^μ (f_{x^μ} is not a tensor, it provides the radius of the space's curvature); these curvature scale factors are needed in the description of curved spaces. The square of the spacetime interval separating two infinitesimally

neighboring events $x^\mu = (ct, x, y, z)$ and $x^\mu + dx^\mu = (ct + cdt,\ x + dx,\ y + dy,\ z + dz)$ is given by

$$\begin{aligned} ds^2 &= ds \cdot ds = \left(\sqrt{f_{x^\mu}}dx^\mu \hat{e}_\mu\right) \cdot \left(\sqrt{f_{x^\nu}}dx^\nu \hat{e}_\nu\right) = \sqrt{f_{x^\mu} f_{x^\nu}}dx^\mu dx^\nu \left(\hat{e}_\mu \cdot \hat{e}_\nu\right) \\ &= g_{\mu\nu}dx^\mu dx^\nu, \qquad \mu,\ \nu = 0, 1, 2, 3, \end{aligned} \tag{D.49}$$

where $g_{\mu\nu}$ is the metric

$$g_{\mu\nu} = \sqrt{f_{x^\mu} f_{x^\nu}}\left(\hat{e}_\mu \cdot \hat{e}_\nu\right). \tag{D.50}$$

Note that we have used Einstein's summation convention in equations (D.48)–(D.49) where implicit summations over μ and ν are implied; in (D.50), however, no summation is implied over μ and ν on the right-hand side.

Assuming that the basis vectors are orthonormal, $\hat{e}_\mu \cdot \hat{e}_\nu = \delta_{\mu\nu}$, the metric (D.50) becomes

$$g_{\mu\nu} = \sqrt{f_{x^\mu} f_{x^\nu}}\delta_{\mu\nu}. \tag{D.51}$$

Hence, we can rewrite (D.49) as follows

$$\begin{aligned} ds^2 &= g_{\mu\nu}dx^\mu dx^\nu = \sqrt{f_{x^\mu} f_{x^\nu}}dx^\mu dx^\nu \delta_{\mu\nu} \\ &\equiv f_{x^\mu}\left(dx^\mu\right)^2 = f_{x^0}\left(dx^0\right)^2 + f_{x^1}\left(dx^1\right)^2 + f_{x^2}\left(dx^2\right)^2 + f_{x^3}\left(dx^3\right)^2. \end{aligned} \tag{D.52}$$

As discussed above, the spacetime metric shows us how to measure the distance separating events.

Metric of the Flat 4-Dimensional Spacetime in Cartesian Coordinates

In the case of a *flat* 4-dimensional Minkowski spacetime where gravity is nonexistent or very weak, the scale factors f_{x^μ} are given by:

$$f_{x^0} = 1, \qquad f_{x^1} = f_{x^2} = f_{x^3} = -1. \tag{D.53}$$

Using these values, we can write the metric elements (D.51) as

$$g_{00} = 1, \qquad g_{11} = g_{22} = g_{33} = -1, \quad \text{and} \quad g_{\mu\nu} = 0 \quad \text{for} \quad \mu \neq \nu. \tag{D.54}$$

Hence, we can write the metric g in matrix as follows

$$[g] = (g_{\mu\nu}) = \begin{pmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{D.55}$$

This is known as the *flat spacetime metric*, also called the *Minkowski metric* or *Minkowski tensor*. Note: As mentioned above, in this text, we use $[g]$ or $(g_{\mu\nu})$ to denote the matrix of the Minkowski metric, while $g_{\mu\nu}$ represents the $\mu\nu$ element of the metric; the same thing applies to the 4-dimensional Kronecker delta function where $[\delta]$ or $(\delta_{\mu\nu})$ denote the matrix and $\delta_{\mu\nu}$ the $\mu\nu$ element of the δ–matrix.

Inner Product in the Flat 4-Dimensional Spacetime

Due to the three minus signs in the metric (D.55), and taking a new basis $(\mathcal{E}_0,\ \mathcal{E}_1,\ \mathcal{E}_2,\ \mathcal{E}_3)$, which

differs slightly[4] from $(\hat{e}_0, \hat{e}_1, \hat{e}_2, \hat{e}_3)$, the inner product between the basis vectors is no longer given by the Euclidean space relation $\hat{e}_\mu \cdot \hat{e}_\nu = \delta_{\mu\nu}$. Combining (D.50), (D.53), and (D.54), we have

$$\hat{e}_\mu \cdot \hat{e}_\nu = g_{\mu\nu} = \begin{cases} 1, & \mu = \nu = 0 \\ -1, & \mu = \nu = 1,2,3 \\ 0, & \mu \neq \nu \end{cases}, \qquad \mu, \nu = 0,1,2,3 \tag{D.56}$$

Hence, we can write the inner product between any two 4-vectors $A^\mu = (A^0, \vec{A}) = (A^0, A^1, A^2, A^3)$ and $B^\mu = (B^0, \vec{B}) = (B^0, B^1, B^2, B^3)$ as follows

$$A \cdot B = \left(\sum_{\mu=0}^{3} A^\mu \hat{e}_\mu\right) \cdot \left(\sum_{\nu=0}^{3} A^\nu \hat{e}_\nu\right) = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} A^\mu B^\nu \left(\hat{e}_\mu \cdot \hat{e}_\nu\right) = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} g_{\mu\nu} A^\mu B^\nu \equiv g_{\mu\nu} A^\mu B^\nu, \tag{D.57}$$

where an implicit summation over μ and ν is implied in the last term. inserting the metric components (D.54) into (D.57), we obtain or as follows

$$\boxed{A \cdot B = g_{\mu\nu} A^\mu B^\nu = A^0 B^0 - \left(A^1 B^1 + A^2 B^2 + A^3 B^3\right) \equiv A^0 B^0 - \vec{A} \cdot \vec{B}.} \tag{D.58}$$

We could have obtained this result directly by simply inserting (D.53) into (D.52); this leads to

$$ds^2 = dx^\mu \cdot dx^\mu = g_{00} c^2 dt^2 + g_{11} dx^2 + g_{22} dy^2 + g_{33} dz^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2, \tag{D.59}$$

where we have used the fact that the components of the infinitesimal displacement vector $dx^\mu = (dx^0, dx^1, dx^2, dx^3)$ are given, in Cartesian coordinates, by $dx^0 = cdt$, $dx^1 = dx$, $dx^2 = dy$, $dx^3 = dz$. Applying (D.59) to the inner product between A^μ and B^μ, we recover (D.58) at once. A particularly interesting application of this relation is to calculate the inner product between two 4-vectors x^μ:

$$x \cdot x = (ct)^2 - (x^2 + y^2 + z^2) = c^2 t^2 - \vec{r}^{\,2}, \tag{D.60}$$

where the first term ct is the distance traveled by light in the allotted time pertaining to the event under study; the second term $x^2 + y^2 + z^2$ represents the square of the distance from the origin to the location where the event occurs. Note that, while the square of a 3-vector $\vec{r}^{\,2} = x^2 + y^2 + z^2$ is always positive in the Euclidean space, the square of a 4-vector is not necessarily positive in the Minkowski spacetime due to the presence of the minus sign in (D.60).

[4]The difference between $(\hat{e}_0, \hat{e}_1, \hat{e}_2, \hat{e}_3)$ and $(\hat{e}_0, \hat{e}_1, \hat{e}_2, \hat{e}_3)$ is in the definition of their respective inner products: while $\hat{e}_\mu \cdot \hat{e}_\nu = \delta_{\mu\nu}$, the other one is given by $\hat{e}_\mu \cdot \hat{e}_\nu = g_{\mu\nu}$.

Appendix E

The Relativistic Notation and Four Vectors

This appendix deals with the Lorentz transformation and the four-vector notation.

E.1 Introduction

Since Chapter 12 is about relativistic quantum mechanics, we would need to offer an outline of the most pertinent mathematical concepts and notations so as to be consistent with special relativity and also to make the formalism easy to follow. The theory of special relativity rests on two postulates

- **Principle of Relativity:** *The laws of physics are the same in all inertial (nonaccelerating) reference frames;* there is no absolute frame of reference in the universe and no absolute state of rest, only relative positions and velocities between observers or objects are measured, hence the term relativity.

- **Constancy of the Speed of Light in Vacuum:** *The Speed of light in vacuum is a universal constant;* it has the same value c in all inertial frames of reference, irrespective of the relative motion of the source and observer.

In essence, special relativity describes the relationship between the results of measurements (observations) made by observers located in different inertial frames in the absence of gravity; if gravity is considered, one has to invoke the theory of general relativity.

One of the biggest problems of the Schrödinger equation is its incompatibility with the theory of special relativity due to its inability to treat the time and space coordinates symmetrically; it treats them as totally independent from each other and asymmetrically (the Schrödinger equation is first order in the time derivative and second order in the space derivatives). In relativity, however, and as we will see below, time and space are treated on equal footing; they are inseparable coordinates.

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

E.2 Lorentz transformation

The *Lorentz transformation* deals with one central issue: How to connect the coordinates of an event that are recorded by one observer to the coordinates of the same event recorded by another observer. The standard approach to achieve this aim is to begin by considering two inertial frames[1] F and F' where F' is in uniform motion along, for instance, the x-direction with constant velocity $\vec{v}$ relative to F as shown in Figure E.1. To connect the coordinates of the same event in the two inertial frames, let us denote the spacetime coordinates of an event by (ct, x, y, z) in F and by (ct', x', y', z') in F'. Now, the problem boils down to finding the mathematical relations that connect (ct', x', y', z') to (ct, x, y, z). Taking F and F' to coincide at $t = t' = 0$, we can write the Lorentz transformation as follows:

$$x' = \gamma(x - c\beta t), \qquad y' = y, \qquad z' = z, \qquad ct' = \gamma(ct - \beta x), \tag{E.1}$$

where $\beta = v/c$ and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ is the *Lorentz factor*. We can rewrite (E.1) as

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma(ct - \beta x) \\ \gamma(x - c\beta t) \\ y \\ z \end{pmatrix}. \tag{E.2}$$

This is also known as a *boost transformation* along the x-axis. We can easily show that the inverse transformation from F' to F can be obtained by simply changing $\vec{v} \longrightarrow -\vec{v}$ (or $\beta \longrightarrow -\beta$):

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma(ct' + \beta x') \\ \gamma(x' + c\beta t) \\ y' \\ z' \end{pmatrix}. \tag{E.3}$$

Remarks

- Since classical physics deals with objects that move with speeds that are too small compared to the speed of light, and in the $c \longrightarrow \infty$ limit (and hence $\beta \longrightarrow 0$), the Lorentz transformation (E.1) reduces, as expected, to the *Galilean transformation* of Newtonian mechanics:

  $$x' = x - vt, \qquad y' = y, \qquad z' = z, \qquad t' = t, \tag{E.4}$$

  which can be written in a matrix form as follows

  $$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}. \tag{E.5}$$

[1] As defined in classical physics and special relativity, frames that have no acceleration relative to one another are known as inertial frames.

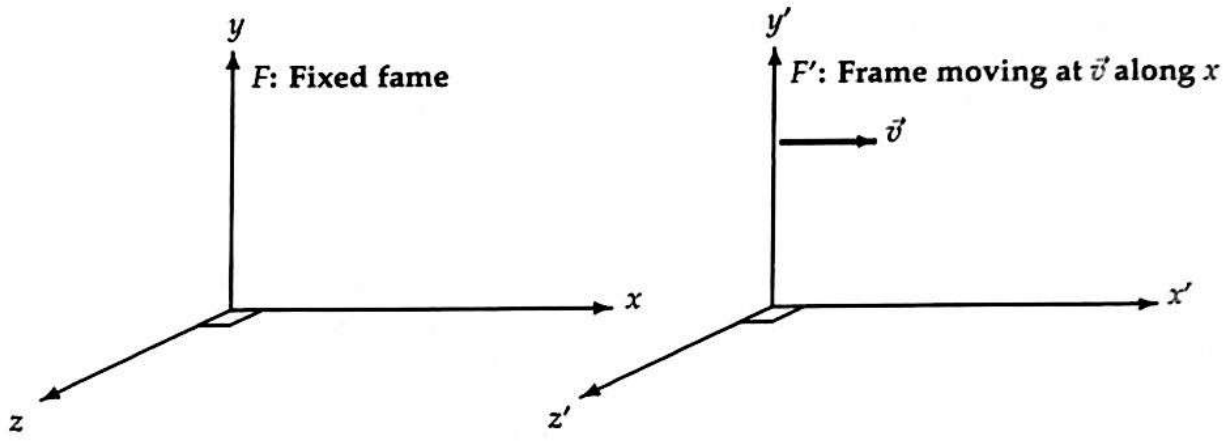


Figure E.1: Inertial frames F and F': F is fixed while F' is moving at constant velocity $\vec{v}$ along the positive x–axis

- Introducing the *rapidity*[2] parameter ϕ, which is defined by $\tanh\phi = \beta = v/c$, we can rewrite the Lorentz transformation (E.2) in the following suggestive form:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}, \tag{E.6}$$

where we have used the facts that

$$\gamma = \frac{1}{\sqrt{1-\tanh^2\phi}} = \frac{\cosh\phi}{\sqrt{\cosh^2\phi - \sinh^2\phi}} = \cosh\phi, \tag{E.7}$$

$$\beta\gamma = \tanh\phi\cosh\phi = \sinh\phi \tag{E.8}$$

Since $\cosh\phi = \cos(i\phi)$ and $i\sinh\phi = \sin(i\phi)$, we can view the Lorentz transformation (E.6) as a hyperbolic rotation through the hyperbolic angle ϕ (the rapidity) in the Minkowski 4-dimensional spacetime, which we are going to consider in Subsection E.3.

- The Lorentz transformation has an important property: it conserves the length of 4-vectors. For instance, and as shown in a solved problem in Chapter 12, we can easily show that the Lorentz transformation (E.2) conserves the length of 4-displacements

$$\boxed{(ct')^2 - (x'^2 + y'^2 + z'^2) = (ct)^2 - (x^2 + y^2 + z^2).} \tag{E.9}$$

[2] *Rapidity* is of special importance in high energy physics where many particles may be produced from a collision between two particles; it is defined as the hyperbolic angle between the momentum of the produced particle and the beam axis. The main reason for using the rapidity instead of the polar angle is that the difference between the rapidities of two particles is invariant with respect to Lorentz boosts along the beam axis; this is why the rapidity is a crucial parameter in accelerator physics.

E.3 Minkowski Spacetime

The motion of an object can only be described relative to something else; for instance, relative to other observers, or objects, or to a frame of reference. The frames of reference and spaces used in relativity and in classical physics are different.

In classical physics, we use the standard 3-dimensional Euclidean space to describe the motion of an object. For instance, when utilizing a Cartesian coordinate system, the position of an object is described by means of its spatial coordinates x, y, z; its position vector is given in terms of the basis vectors $\hat{i}$, $\hat{j}$, $\hat{k}$ by $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$, which we will be denoting by $\vec{r} \equiv (x, y, z)$. The time evolution of the object's position vector is described by $\vec{r}(t) = (x(t), y(t), z(t))$, where t is an *absolute parameter* which is independent of the spatial coordinates x, y, z.

In special relativity, however, one does not deal with the 3-dimensional Euclidean space; one uses a different space – a space in which the rules of special relativity hold. In this context, Minkowski (one of Einstein's professors at the Zurich Polytechnic, now known as the Swiss Federal Institute of Technology (ETH Zurich)) has shown in 1908 that the theory of relativity can be best described within the framework of a *four dimensional space*, known ever since then as the *Minkowski spacetime*[3]. Unlike the Euclidean space, which deals with the spatial coordinates only, the Minkowski spacetime includes time in addition to the spatial coordinates in a way that the 3 spatial coordinates and the 1 time component are not separate entities, they are inseparably connected. As a result, the Minkowski spacetime allows us to deal with processes or events that take place over time illustrating effects such as length contraction, time dilation, and simultaneity. We should also note that, in special relativity and unlike classical physics, time is not an absolute parameter. However, the time resulting from the reading of a clock carried by the particle (i.e., the time in the particle's own rest frame) is an invariant parameter[4] that all observers agree on, since it is a scalar under Lorentz transformations; it is known as the *proper time*.

A point in the Minknowski spacetime is characterised by what happens or takes place there; we will call this an *"event"*; an event is something that occurs independently of the fame of reference used to describe it. For example, two particles colliding at a certain point in space is an event; this event takes place at a specific place (in the 3-dimensional space) and a unique time seen by an observer in a given inertial reference frame. As a result, an event in Minkowski spacetime is defined or indexed by *four components*: 3 spatial coordinates and 1 time component. This topic becomes clearer when we deal next with the 4-vector notation. We should mention that one of the most important usefulness of 4-vectors is the fact that they lend themselves quite easily to transform their components between inertial frames using the Lorentz transformation as we will see in Section E.2.

E.4 The 4-Vector Notation

Four Vectors

As mentioned above, *an event in the Minkowski spacetime is represented by a set of four coordinates, one for time and three for space*: (ct, x, y, z). Note that we have multiplied the time with the speed

[3] Hermann Minkowski, *Space and Time*, *Physikalische Zeitschrift*, **10**, pp. 75-88 (1908).

[4] We should note that time invariance is not synonymous to being constant; it simply means the time span between two events is the same in all inertial frames.

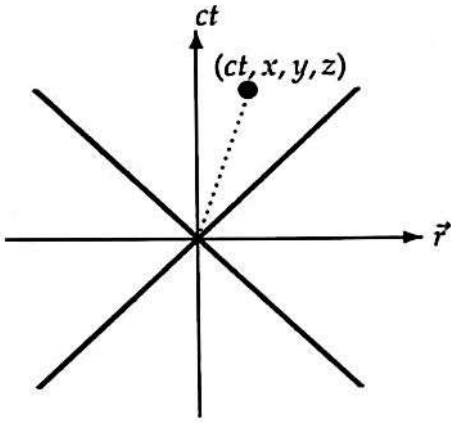


Figure E.2: Spacetime diagram where the location of an event in the Minkowski spacetime is represented by a set of four coordinates (ct, x, y, z) where $\vec{r}$ stands for $(x,\ y,\ z)$.

of light c in the first component (i.e., ct) to ensure that all such four components have the same physical dimensions of length; we measure the time elapsed in terms of how far light travels in a given time. A pictorial view of an event in the Minkowski spacetime is depicted in Figure E.2 (where $\vec{r}$ stands for a maximum of two spacial coordinates due to the drawing limitations inherent in the spacetime diagram model as also illustrated in Figure E.3).

Denoting the time-like component by $x^0 = ct$ (i.e., a scalar) and the 3 space-like components by $x^1 = x,\ x^2 = y,\ x^3 = z$ (i.e., a 3-vector), we can define a *4-vector* x^μ that unifies time and space components together in the following simplifying form[5]:

$$x^\mu = \left(x^0,\ x^1,\ x^2,\ x^3\right) = (ct,\ x,\ y,\ z) \equiv (ct,\ \vec{r}), \qquad \mu = 0,\ 1,\ 2,\ 3; \tag{E.10}$$

x^μ is also known as the event's 4-displacement vector in spacetime; it is used in relativity to describe the motion of an object. In spacetime, the object's "trajectory", which consists of a succession of events where each event is specified by x^μ, is known as the *world line* of the object where time is no longer an absolute parameter, for it gives the reading of a clock carried by the object under study; i.e., the time in the particle's own rest frame and, hence, is independent of coordinates. Time t is an invariant parameter as it is not specific to any coordinate system; it is called the proper time.

By analogy to the 4-displacement vector (E.10) where space and time are inseparable, energy and momentum are also treated on equal footing. As such, we can introduce the 4-momentum p^μ (the counterpart of the 3-momentum vector, $\vec{p}$, of the Euclidean space) of a particle of energy E and momentum $\vec{p}$ as follows

$$p^\mu = \left(p^0,\ p^1,\ p^2,\ p^3\right) = \left(\frac{E}{c},\ \vec{p}\right). \tag{E.11}$$

We can also introduce the 4-current density j^μ, the analog of the 3-dimensional current density

[5]**Important Note:** The superscripts 0, 1, 2, 3 appearing in Equation (E.10) are not exponents; they are used to label the various components. For instance, the superscript "3" in "x^3" refers to coordinate 3 (the z-component in this case), not the cube power of "x", and so on. The meaning of the superscripts will become clearer in Subsection E.5 when we deal with contravariant and covariant notations.

vector $\vec{j}(\vec{r}, t)$ and the charge density $\rho(\vec{r}, t)$, as follows

$$j^\mu = (c\rho, \vec{j}) = \left(c\rho(\vec{r}, t), \vec{j}(\vec{r}, t)\right), \tag{E.12}$$

the 4-vector potential A^μ, which describes the coupling of a particle of charge q (and electric potential U) to an electromagnetic field in terms of the 3-dimensional vector potential $\vec{A}$, in SI units as

$$A^\mu = (A^0, \vec{A}) = \left(\frac{U}{c}, \vec{A}\right) \tag{E.13}$$

and the 4-wave vector k^μ, the counterpart of the 3-dimensional wave vector $\vec{k} = \vec{p}/\hbar$, as follows:

$$k^\mu = \frac{1}{\hbar}p^\mu = \left(\frac{E}{\hbar c}, \frac{\vec{p}}{\hbar}\right) = \left(\frac{\omega}{c}, \vec{k}\right), \tag{E.14}$$

where we have used the fact that $E = \hbar\omega$. Notice that, in equations (E.11)–(E.14), we have either divided or multiplied the first components by the speed of light c so as to ensure that the four components of p^μ, j^μ, A^μ, and k^μ have the same physical dimensions, each separately.

Remarks

For dimensional consistency, we have made sure that the time-like and space-like components have the same physical dimensions in all the 4–vectors considered above. The obvious example is that the time-like term E/c in p^μ has the dimensions a momentum (since Energy / speed ≡ momentum) very much like the space-like component $\vec{p}$. Another, perhaps less obvious, example is A^μ where the physical dimensions[6] of U/c and $\vec{A}$ are

$$[U/c] = \mathrm{MLT^{-1}Q^{-1}}, \qquad [\vec{A}] = \mathrm{MLT^{-1}Q^{-1}}, \tag{E.15}$$

where M = mass, L = Length, T = Time, and Q = Charge. Hence, the physical dimensions of qU/c and $q\vec{A}$ are those of a linear momentum $\vec{p}$:

$$[qU/c] = \mathrm{MLT^{-1}}, \qquad [q\vec{A}] = \mathrm{MLT^{-1}}. \tag{E.16}$$

As presented in Appendix D, we can greatly simplify the 4-vector formalism by adopting a more general notation, the *index notation* where a 4-vector A^μ is represented by its four components as follows[7]

$$A^\mu = (A^0, \vec{A}) = (A^0, A^1, A^2, A^3), \qquad \text{where} \qquad A^1 = A_x,\ A^2 = A_y,\ A^3 = A_z; \tag{E.17}$$

$\mu = 0$ represents the 'time component' and $\mu = 1, 2, 3$ the 'space components.' As shown in Appendix D, the inner product (or 'dot' product) $A \cdot B$ between any two 4-vectors $A^\mu = (A^0, A^1, A^2, A^3)$ and $B^\mu = (B^0, B^1, B^2, B^3)$ is given by

$$A \cdot B = A^0B^0 - \left(A^1B^1 + A^2B^2 + A^3B^3\right) \equiv A^0B^0 - \vec{A} \cdot \vec{B}. \tag{E.18}$$

---

[6] Since $A^\mu = (A^0, \vec{A}) \equiv (U/c, \vec{A})$, where U is the electric potential (i.e., voltage) and c is the speed of light, we can infer the physical dimensions of the electric potential U form the expression of energy: [Electric Potential]= [Energy] / [Charge]. Now, since the physical dimensions of energy are $[E]=\mathrm{ML^2T^{-2}}$, we have $[U]=\mathrm{ML^2T^{-2}Q^{-1}}$ and, hence $[U/c]=\mathrm{ML^2T^{-2}Q^{-1}} / (\mathrm{LT^{-1}}) = \mathrm{MLT^{-1}Q^{-1}}$.

[7] The 4-vector A^μ we are using here is general and should not be confused with the 4-vector potential of Equation (E.13)

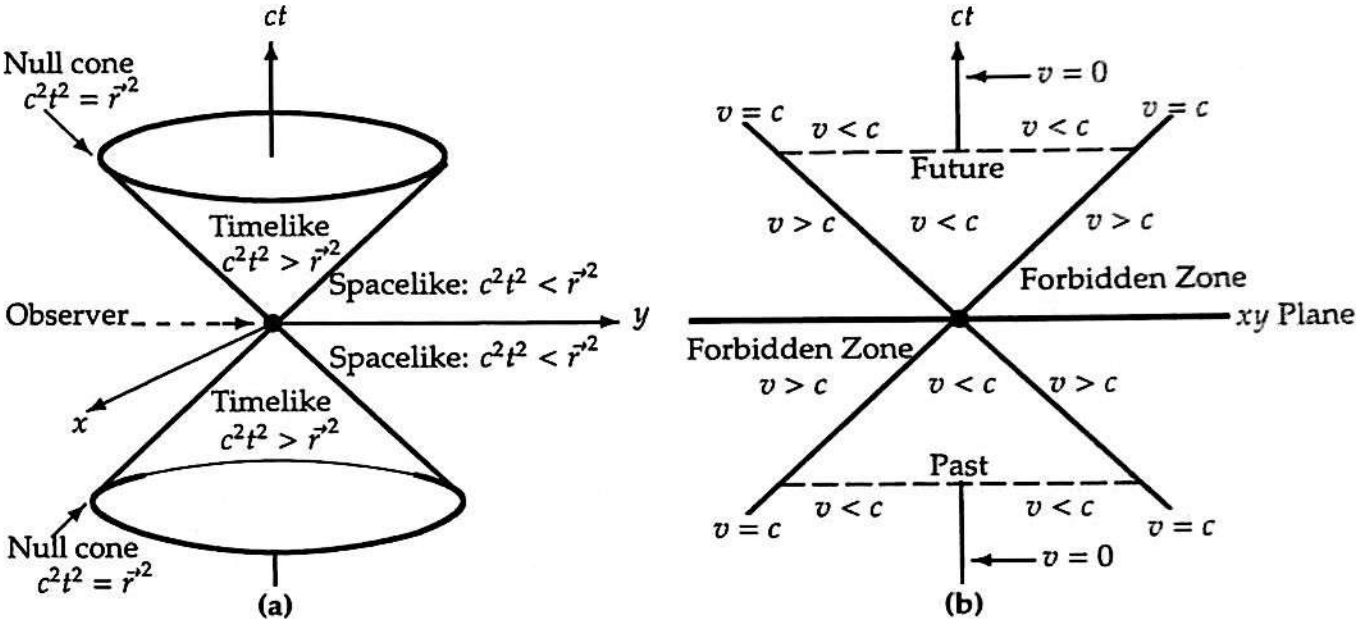


Figure E.3: **(a)**: Conventional three-dimensional light cone diagram with 2-space xy-dimensions (the z-dimension is suppressed) and one time dimension.**(b)**: Two-dimensional representation of the light cone diagram of an event (with speed v).

Applying this relation to (E.10), we can write

$$x^2 \equiv x \cdot x = (ct)^2 - (x^2 + y^2 + z^2) = c^2t^2 - \vec{r}^2, \tag{E.19}$$

where the first term ct is the distance traveled by light in the allotted time pertaining to the event under study; the second term $x^2 + y^2 + z^2$ represents the square of the distance from the origin to the location where the event occurs.

Note that, while the square of a 3-vector $\vec{r}^2 = x^2 + y^2 + z^2$ is always positive in the Euclidean space, *"the square[8] of a 4-vector" is not necessarily positive in the Minkowski spacetime due to the presence of the minus sign in* (E.19); this means that a 4-vector x^μ can be one of these three types: either $x^2 > 0$ or $x^2 = 0$ or $x^2 < 0$. This negative sign will have profound implications as will be seen right below.

Pictorial view of events in the Minkowski spacetime: Light Cone

Minkowski has pictorially represented spacetime as a *light cone* which defines a boundary between past and future accessible events and consists of three zones corresponding to $x^2 > 0$, $x^2 = 0$, $x^2 < 0$. As Figure E.3-(a) shows, the light cone is a 3-dimensional surface in the 4-dimensional spacetime. Events in spacetime may be characterized according to whether they are taking place inside of, outside of, or on the light cone; this divides spacetime into three zones about x^μ. So, a 4-vector x^μ is called *timelike, lightlike,* or *spacelike* depending on

[8]This should not be confused with the ordinary square of a 3-vector $\vec{r}^2 = x^2 + y^2 + z^2$ in the Euclidean space.

whether its magnitude square $x^2 = c^2t^2 - \vec{r}^2$ is positive, zero, or negative, respectively; that is, the event's 4-vector x^μ is said to be

- **Timelike** if $x^2 > 0$ or $c^2t^2 > r^2$ (or $v < c$; this is the causal zone),
- **Lightlike** (or null) if $x^2 = 0$ or $c^2t^2 = r^2$ (or $v = c$, this is the null or light cone; this represents a spherical wave front of light propagating outwards from the cone's origin and reaching the radius $r = ct = \sqrt{x^2 + y^2 + z^2}$), and
- **Spacelike** if $x^2 < 0$ or $c^2t^2 < r^2$ (or $v > c$, where v is the speed of the event; this is the forbidden or non-causal zone).

From the theory of special relativity and as illustrated in Figure E.3-(b), we may recall the fact that, while space-like events cannot be connected by real physical processes, *time-like events can be connected by causal, real physical signals*. The light-like events pertain only to those situations where objects move at the speed of light (i.e., $v = c$); the $v = c$ surface separates the causal from the forbidden zones as depicted in Figure E.3-(b). While the separation between time-like events is real and non-zero, the separation between space-like events is imaginary. The separation between light-like is zero; electromagnetic radiation can travel only on light-like trajectories. It helps to note that space-like events are impossible for they can, at least theoretically, travel faster than light; they would seem to be traveling backwards in time as seen by some observers.

E.5 Contravariant and Covariant notations

The underlying reason for introducing the covariant notation is its usefulness in demonstrating the relativistic covariance (or invariance) of various quantities and relations; relativistic covariance is easier to prove when the covariant notation is used. Equations that are consistent with special relativity can be written in terms of 4-vectors and tensors (i.e., in covariant form); it then becomes easier to see if they hold the same form in all inertial frames. For instance, and as will be seen below (see equation (E.76)), the covariant notation will make it easy to demonstrate that inner products of 4-vectors are invariant under Lorentz transformations.

As mentioned above, unlike the positive definite nature of the magnitude square of 3-vectors $r^2 = x^2 + y^2 + z^2$, *the magnitude square of 4-vectors is not necessarily positive definite* due to the *minus sign* associated with the spatial components as shown in (E.19). To obtain a Lorentz invariant form of the magnitude square and to properly keep track of the minus sign associated with the spatial components, we need to introduce the *contravariant and covariant* 4-vector notations.

The **upper index** vector x^μ introduced in (E.10) is known as the **contravariant** form of the 4-vector x; its **covariant** form, which is denoted by the **lower index** vector x_μ, is obtained by merely changing the sign of the spatial components in (E.10):

$$\boxed{\text{Contravariant 4-vector:} \quad x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) \equiv (x^0, \vec{r}),} \tag{E.20}$$

$$\boxed{\text{Covariant 4-vector:} \quad x_\mu = (x_0, x_1, x_2, x_3) = (ct, -x, -y, -z) \equiv (x^0, -\vec{r}).} \tag{E.21}$$

So, contravariant and covariant 4-vectors x^μ and x_μ have the same time components but their spatial components differ by a minus sign; covariant vectors are just spatially inverted contravariant vectors:

$$x^0 = x_0, \qquad x^j = -x_j, \qquad \text{where} \qquad j = 1, 2, 3. \tag{E.22}$$

By analogy to (E.20) and (E.21) and using (E.11)–(E.14), we now can introduce the contravariant and covariant 4-vectors corresponding to the momentum, current density, electromagnetic 4-vector potential, and wave vector as follows:

$$p^\mu = \left(p^0, p^1, p^2, p^3\right) = \left(\frac{E}{c}, \vec{p}\right), \qquad p_\mu = \left(p^0, -p^1, -p^2, -p^3\right) = \left(\frac{E}{c}, -\vec{p}\right). \tag{E.23}$$

$$j^\mu = (c\rho, \vec{j}), \qquad j_\mu = (c\rho, -\vec{j}), \tag{E.24}$$

$$A^\mu = (A^0, \vec{A}) = \left(\frac{U}{c}, \vec{A}\right), \qquad A_\mu = (A^0, -\vec{A}) = \left(\frac{U}{c}, -\vec{A}\right), \tag{E.25}$$

$$k^\mu = \frac{1}{\hbar}p^\mu = \left(\frac{E}{\hbar c}, \frac{\vec{p}}{\hbar}\right) = \left(\frac{\omega}{c}, \vec{k}\right), \qquad k_\mu = \frac{1}{\hbar}p_\mu = \left(\frac{E}{\hbar c}, -\frac{\vec{p}}{\hbar}\right) = \left(\frac{\omega}{c}, -\vec{k}\right). \tag{E.26}$$

Remarks

For dimensional consistency, we have made sure that the time-like and space-like components have the same physical dimensions in all the 4–vectors considered above. The obvious example is that the time-like term E/c in p^μ has the dimensions a momentum (since Energy / speed ≡ momentum) very much like the space-like component $\vec{p}$. Another, perhaps less obvious, example is A^μ where the physical dimensions[9] of U/c and $\vec{A}$ are

$$[U/c] = \mathrm{MLT^{-1}Q^{-1}}, \qquad [\vec{A}] = \mathrm{MLT^{-1}Q^{-1}}, \tag{E.27}$$

where M = mass, L = Length, T = Time, and Q = Charge. Hence, the physical dimensions of qU/c and $q\vec{A}$ are those of a linear momentum $\vec{p}$:

$$[qU/c] = \mathrm{MLT^{-1}}, \qquad [q\vec{A}] = \mathrm{MLT^{-1}}. \tag{E.28}$$

Contravariant-covariant notations for the 4-momentum and 4-gradient operators

Inserting the operators $E \longrightarrow i\hbar\partial/\partial t$ and $\vec{p} \longrightarrow -i\hbar\vec{\nabla}$ into (E.23), we can write the contravariant and covariant 4-momentum operator[10], respectively, as follows:

$$\boxed{\hat{p}^\mu = i\hbar\partial^\mu = i\hbar\frac{\partial}{\partial x_\mu} = \left(\frac{i\hbar}{c}\frac{\partial}{\partial t}, -i\hbar\vec{\nabla}\right), \qquad \hat{p}_\mu = i\hbar\partial_\mu = i\hbar\frac{\partial}{\partial x^\mu} = \left(\frac{i\hbar}{c}\frac{\partial}{\partial t}, i\hbar\vec{\nabla}\right),} \tag{E.29}$$

[9]Since $A^\mu = (A^0, \vec{A}) \equiv (U/c, \vec{A})$, where U is the electric potential (i.e., voltage) and c is the speed of light, we can infer the physical dimensions of the electric potential U form the expression of energy: [Electric Potential]= [Energy] / [Charge]. Now, since the physical dimensions of energy are $[E]=\mathrm{ML^2T^{-2}}$, we have $[U]=\mathrm{ML^2T^{-2}Q^{-1}}$ and, hence $[U/c]=\mathrm{ML^2T^{-2}Q^{-1}} / (\mathrm{LT^{-1}}) = \mathrm{MLT^{-1}Q^{-1}}$.

[10]Notice the notational difference between the 4-momentum vector p^μ and the 4-momentum operator $\hat{p}^\mu$ is achieved by the presence of the "hat" sign in the operator as described in Section 2.4 of Chapter 2.

where ∂^μ and ∂_μ represent, respectively, the contravariant and covariant forms of the differential operator

$$\boxed{\text{Contravariant 4-gradient: } \partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\partial^0, -\vec{\nabla}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\vec{\nabla}\right) \equiv \left(\frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right),} \tag{E.30}$$

which can also be written as

$$\partial^\mu = \left(\partial^0, \partial^1, \partial^2, \partial^3\right) = \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial x_0}, \frac{\partial}{\partial x_1}, , \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right), \tag{E.31}$$

and

$$\boxed{\text{Covariant 4-gradient: } \partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\partial_0, \vec{\nabla}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla}\right) \equiv \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right),} \tag{E.32}$$

which can be written as

$$\partial_\mu = (\partial_0, \partial_1, \partial_2, \partial_3) = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, , \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right). \tag{E.33}$$

Notice the sign differences between the contravariant and covariant forms of the 4-vectors (E.20)–(E.21) and the 4-gradients operators (E.30)–(E.32). This shows that the contravariant gradient operator ∂^μ is the derivative with respect to the covariant vector x_μ (i.e., $\partial^\mu = \partial/\partial x_\mu$) and, conversely, the covariant gradient operator ∂_μ is the derivative with respect to the contravariant vector x^μ (i.e., $\partial_\mu = \partial/\partial x^\mu$).

Connecting contravariant and covariant vectors by the Minkowski metric
Contravariant and covariant 4-vectors can be transformed into each other as follows

$$x_\mu = \sum_{\nu=0}^{3} g_{\mu\nu}x^\nu = g_{\mu 0}x^0 + g_{\mu 1}x^1 + g_{\mu 2}x^2 + g_{\mu 3}x^3, \tag{E.34}$$

where $g_{\mu\nu}$ are the matrix elements of what is known as the *flat spacetime metric*[11] $[g]$, also called the *Minkowski metric* or the *Minkowski tensor*. Note: In this text, we use $[g]$ or $(g_{\mu\nu})$ to denote the matrix of the Minkowski metric, while $g_{\mu\nu}$ represents the $\mu\nu$ element of the metric's matrix; the same thing applies to the 4-dimensional Kronecker delta function where $[\delta]$ or $(\delta_{\mu\nu})$ denote the matrix and $\delta_{\mu\nu}$ the $\mu\nu$ element of the δ–matrix.. As shown in Appendix D (Section D.3), only the diagonal elements of $[g]$ are not equal to zero (i.e., $g_{\mu\nu} = 0$ for $\mu \neq \nu$, $g_{00} = 1$, and $g_{11} = g_{22} = g_{33} = -1$):

$$[g] = \left(g_{\mu\nu}\right) = \begin{pmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{E.35}$$

[11] In this chapter, we will be working in flat space (no curvature) only.

Note: a heuristic derivation of this Minkowski metric is presented below (see Equation (E.52)) by means of a matrix representation of covariant and contravariant vectors.

We can write (E.34) in a simpler form by using the *Einstein summation convention* which does away with the explicit appearance of the summation symbol:

$$x_\mu = \sum_{\nu=0}^{3} g_{\mu\nu} x^\nu \equiv g_{\mu\nu} x^\nu. \tag{E.36}$$

According to this summation convention, whenever an index variable appears twice in a single term (here ν), summation of that term over all the values of the index is automatically implied (here $\nu = 0, 1, 2, 3$), even though the summation symbol is not explicitly written. This is one of the advantages inherent in the index notation of vectors. For notational brevity, we will be using the Einstein summation convention throughout Chapter 12 and the relevant appendices D, E, F, and G.

Defining the contravariant Minkowski metric in terms of its elements $g^{\mu\nu}$, which represent the $\mu\nu$ element of the inverse matrix $[g]^{-1}$, and taking the inverse of (E.35), we obtain

$$[g]^{-1} = (g^{\mu\nu}) = \begin{pmatrix} g^{00} & g^{01} & g^{02} & g^{03} \\ g^{10} & g^{11} & g^{12} & g^{13} \\ g^{20} & g^{21} & g^{22} & g^{23} \\ g^{30} & g^{31} & g^{32} & g^{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{E.37}$$

Comparing (E.35) and (E.37), we see that the matrix $[g]$ is equal to its inverse[12] $[g]^{-1}$:

$$[g] = [g]^{-1} \quad \text{or} \quad g_{\mu\nu} = g^{\mu\nu}, \quad [g][g]^{-1} = [g]^{-1}[g] = I_4, \quad [g]^2 = I_4, \tag{E.38}$$

where I_4 is the 4-rowed identity matrix

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{E.39}$$

Notice that the Minkowski matrix (E.35) and its inverse (E.37) are symmetric

$$g_{\mu\nu} = g_{\nu\mu}, \qquad g^{\mu\nu} = g^{\nu\mu}. \tag{E.40}$$

We can verify that the product between the contravariant $g^{\mu\nu}$ and covariant $g_{\mu\nu}$ forms of the Minkowski matrix $[g]$ is equal to the 4-dimensional Kroneker delta function[13]:

$$\left(g_{\mu\lambda} g^{\lambda\nu}\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \left(\delta^\nu_\mu\right) = \left(\delta_{\mu\nu}\right) = (\delta^{\mu\nu}) \equiv I_4, \tag{E.41}$$

[12]This is valid only because the spacetime of special relativity is *flat* (no curvature); this is not the case in general, especially when the spacetime is bent by strong gravitational fields. In the theory of general relativity, massive objects warp the spacetime around them; hence, locally, spacetime is *curved* around every object with mass.

[13]Notice the notational difference between $\delta_{\mu\nu}$, which denotes the $\mu\nu$ matrix element, and $\left(\delta_{\mu\nu}\right)$, which represents the 4×4 matrix of the Kroneker delta function.

where $\delta_{\mu\nu} = 0$ for $\mu \neq \nu$ and $\delta_{\mu\mu} = 1$. From these expressions, we can infer the following useful relations

$$\delta_{\mu\nu} = \frac{\partial x^\mu}{\partial x^\nu} = \frac{\partial x_\mu}{\partial x_\nu}, \qquad g_{\mu\nu} = \frac{\partial x^\mu}{\partial x_\nu}, \qquad g_{\mu\nu}g^{\mu\nu} = 4. \tag{E.42}$$

So, the Minkowski metric $g_{\mu\nu}$ and its inverse $g^{\mu\nu}$ are needed to raise and lower the indices of 4-vectors and tensors:

$$\boxed{x_\mu = g_{\mu\nu}x^\nu, \qquad x^\mu = g^{\mu\nu}x_\nu.} \tag{E.43}$$

We see that the Minkowski metric converts contravariant vectors into covariant vectors and vice versa. While $g_{\mu\nu}$ is used to lower indices, $g^{\mu\nu}$ is used to raise indices; $g_{\mu\nu}$ is called the covariant metric, and $g^{\mu\nu}$ the contravariant metric.

Now, since the off-diagonal elements of $g_{\mu\nu}$ and $g^{\mu\nu}$ are zero (i.e., $g_{\mu\nu} = g^{\mu\nu} = 0$ for $\mu \neq \nu$), an insertion of (E.35) into (E.34) and (E.37) into $x^\mu = g^{\mu\nu}x_\nu$ leads to the following respective relations

$$x_0 = g_{00}x^0 = x^0, \quad x_1 = g_{11}x^1 = -x^1, \quad x_2 = g_{22}x^2 = -x^2, \quad x_3 = g_{33}x^3 = -x^3, \tag{E.44}$$

$$x^0 = g^{00}x_0 = x^0, \quad x^1 = g^{11}x_1 = -x_1, \quad x^2 = g^{22}x_2 = -x_2, \quad x^3 = g^{33}x_3 = -x_3. \tag{E.45}$$

Inner product between 4-vectors in the contravariant-covariant notation

Using (E.18), we can write the inner product $A{\cdot}B$ between any two 4-vectors $A^\mu = (A^0, A^1, A^2, A^3)$ and $B^\mu = (B^0, B^1, B^2, B^3)$ in the Minkowski metric (E.35) as

$$A \cdot B = A^\mu B_\mu = A_\mu B^\mu = g_{\mu\nu}A^\mu B^\nu = g^{\mu\nu}A_\mu B_\nu = A^0B^0 - A^iB^i, \tag{E.46}$$

where we have used the Einstein summation convention which implies summations over the repeated indices $\mu, \nu = 0, 1, 2, 3$ and $i = 1, 2, 3$. In general, we will use Greek letters for space-time indices (which take the values 0, 1, 2, 3) and Roman letters for space indices (which take the values 1, 2, 3); that is:

$$A^\mu B_\mu \equiv \sum_{\mu=0}^{3} A^\mu B_\mu, \qquad A^iB^i \equiv \sum_{i=1}^{3} A^iB^i. \tag{E.47}$$

We should note that *inner products of 4-vectors can be taken only between a contravariant and a covariant vector* (e.g., $A^\mu B_\mu$), *not between two covariant or two contravariant vectors*; the quantity $A^\mu B^\mu = A^0B^0 + A^iB^i$ does not represent the inner product between A^μ and B^μ. Note that while $A^\mu B_\mu$ is invariant under Lorentz transformations, the quantity $A^\mu B^\mu$ is not (see Example 12.2 for a proof).

From (E.46), it is self-evident that the inner product of two 4-vectors is independent of their order:

$$A \cdot B = B \cdot A \qquad \longrightarrow \qquad A_\mu B^\mu = B^\mu A_\mu. \tag{E.48}$$

Using the 3-dimensional Laplace operator $\nabla^2 = \partial_i\partial^i = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ along with (E.30), (E.32) and (E.46), we can write the inner product between the gradient operators ∂_μ and ∂^μ as follows

$$\partial_\mu\partial^\mu = \left(\frac{\partial}{\partial x^0}\right)^2 - \nabla^2 = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \equiv \Box. \tag{E.49}$$

This is called the *d'Alembert operator* or the *d'Alembertian* which we have denoted by the symbol $\Box$. The d'Alembertian $\partial_\mu\partial^\mu$ is the 4-dimensional analogue of the 3-dimensional Laplacian operator $\partial_i\partial^i = \nabla^2$.

We can write the action of the 4-gradient operator on a 4-vector B^μ as follows:

$$\partial_\mu B^\mu = \frac{1}{c}\frac{\partial B^0}{\partial t} + \vec{\nabla}\cdot\vec{B}. \tag{E.50}$$

A summary of all the 4-vectors and operators introduced above along with their 3-dimensional counterparts is listed in Chapter 12 (see Table 12.1).

Matrix representations of contravariant and covariant vectors and their inner products
Representing 4-vectors by column matrices, we are going to show how to infer the Minkowski metric (E.35). For this, by writing covariant and contravariant 4-vectors as follows,

$$x_\mu = \begin{pmatrix} ct \\ -x \\ -y \\ -z \end{pmatrix}, \qquad x^\mu = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}, \tag{E.51}$$

we can easily infer the following 4×4 matrix that connects them:

$$x_\mu = \begin{pmatrix} ct \\ -x \\ -y \\ -z \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv \left(g_{\mu\nu}\right)x^\nu \quad\Longrightarrow\quad \left(g_{\mu\nu}\right) = [g] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}; \tag{E.52}$$

this is nothing but the Minkowski metric obtained above, (E.35), which serves to convert covariant into contravariant components. Following the same method, we can obtain the inverse of the Minkowski metric which connects a contravariant 4-vector to its covariant counterpart

$$x^\mu = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} ct \\ -x \\ -y \\ -z \end{pmatrix} \equiv (g^{\mu\nu})x_\nu \quad\Longrightarrow\quad (g^{\mu\nu}) = [g]^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{E.53}$$

Again, this is the inverse of the Minkowski metric that we have seen above (see (E.37)). So, equations (E.52) and (E.53) give the matrix representation of (E.43).

We can also represent the inner product (E.46) in a matrix form as follows:

$$A\cdot B = A^T g B, \tag{E.54}$$

where A^T is the transpose of the 4-vector A (note that the transpose is represented by a row

matrix); we can write (E.54) as

$$\begin{aligned} A^\mu g_{\mu\nu} B^\nu &= \begin{pmatrix} A^0 & A^1 & A^2 & A^3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} B^0 \\ B^1 \\ B^2 \\ B^3 \end{pmatrix} \\ &= \begin{pmatrix} A^0 & -A^1 & -A^2 & -A^3 \end{pmatrix} \begin{pmatrix} B^0 \\ B^1 \\ B^2 \\ B^3 \end{pmatrix} \\ &= A^0 B^0 - \left(A^1 B^1 + A^2 B^2 + A^3 B^3 \right), \end{aligned} \tag{E.55}$$

or in a simpler matrix form as

$$\begin{aligned} A_\mu B^\mu &= \begin{pmatrix} A_0 & A_1 & A_2 & A_3 \end{pmatrix} \begin{pmatrix} B^0 \\ B^1 \\ B^2 \\ B^3 \end{pmatrix} \equiv \begin{pmatrix} A^0 & -A^1 & -A^2 & -A^3 \end{pmatrix} \begin{pmatrix} B^0 \\ B^1 \\ B^2 \\ B^3 \end{pmatrix} \\ &= A^0 B^0 - \left(A^1 B^1 + A^2 B^2 + A^3 B^3 \right). \end{aligned} \tag{E.56}$$

Similarly, we can represent the magnitude square (E.19) in the following matrix form:

$$\begin{aligned} x_\mu x^\mu &= \begin{pmatrix} x_0 & x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} ct & -x & -y & -z \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \\ &= (ct)^2 - \left(x^2 + y^2 + z^2 \right). \end{aligned} \tag{E.57}$$

As expected, equations (E.55) and (E.56 show that inner products are scalars. However, covariant and contravariant vectors are represented by row and column matrices, respectively.

E.6 Lorentz transformation in the contravariant-covariant form

We can write the Lorentz transformation (E.2) in a matrix form as follows

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}, \tag{E.58}$$

or recast it in the following more compact form by using the 4-vector notation[14]

$$x' = [\Lambda]\, x \qquad \text{or} \qquad x'^\mu = \sum_{\nu=0}^{3} \Lambda^\mu{}_\nu x^\nu \equiv \Lambda^\mu{}_\nu x^\nu, \tag{E.59}$$

[14]Here x and x' are column 4-vectors (not row vectors), $[\Lambda]$ is a 4×4 matrix, x'^μ is the μth component of x', x^ν is the νth component of x, and so on.

where all the elements, $\Lambda^{\mu}{}_{\nu}$ (where μ is the row index and ν is the column index), of the Lorentz transformation matrix $[\Lambda]$ are zero, except

$$\Lambda^{0}{}_{0} = \Lambda^{1}{}_{1} = \gamma, \quad \Lambda^{0}{}_{1} = \Lambda^{1}{}_{0} = -\beta\gamma, \quad \Lambda^{2}{}_{2} = \Lambda^{3}{}_{3} = 1; \tag{E.60}$$

hence, we have

$$[\Lambda] \equiv \left(\Lambda^{\mu}{}_{\nu}\right) = \begin{pmatrix} \Lambda^{0}{}_{0} & \Lambda^{0}{}_{1} & \Lambda^{0}{}_{2} & \Lambda^{0}{}_{3} \\ \Lambda^{1}{}_{0} & \Lambda^{1}{}_{1} & \Lambda^{1}{}_{2} & \Lambda^{1}{}_{3} \\ \Lambda^{2}{}_{0} & \Lambda^{2}{}_{1} & \Lambda^{2}{}_{2} & \Lambda^{2}{}_{3} \\ \Lambda^{3}{}_{0} & \Lambda^{3}{}_{1} & \Lambda^{3}{}_{2} & \Lambda^{3}{}_{3} \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{E.61}$$

This transformation is known as the *Lorentz boost* along the x–axis.

The *inverse* of (E.59), which is obtained by simply multiplying both sides of (E.59) by $[\Lambda]^{-1}$, connects (x^0, x^1, x^2, x^3) to (x'^0, x'^1, x'^2, x'^3):

$$x = [\Lambda]^{-1}x' \qquad \text{or} \qquad x^{\mu} = \Lambda_{\nu}{}^{\mu}x'^{\nu} \equiv \left(\Lambda^{-1}\right)^{\mu}{}_{\nu}x'^{\nu}, \tag{E.62}$$

where $\Lambda_{\nu}{}^{\mu} \equiv \left(\Lambda^{-1}\right)^{\mu}{}_{\nu}$ is the inverse of $\Lambda^{\mu}{}_{\nu}$:

$$[\Lambda]^{-1}[\Lambda] = I_4 \qquad \Longrightarrow \qquad \left(\Lambda^{-1}\right)^{\mu}{}_{\alpha}\Lambda_{\nu}{}^{\alpha} \equiv \Lambda_{\alpha}{}^{\mu}\Lambda^{\alpha}{}_{\nu} = \delta_{\mu\nu}; \tag{E.63}$$

that is, the inverse of (E.61) is given by:

$$\begin{aligned} [\Lambda]^{-1} \equiv \left(\left(\Lambda^{-1}\right)^{\mu}{}_{\nu}\right) &= \begin{pmatrix} \Lambda^{0}{}_{0} & -\Lambda^{1}{}_{0} & -\Lambda^{2}{}_{0} & -\Lambda^{3}{}_{0} \\ -\Lambda^{0}{}_{1} & \Lambda^{1}{}_{1} & \Lambda^{2}{}_{1} & \Lambda^{3}{}_{1} \\ -\Lambda^{0}{}_{2} & \Lambda^{1}{}_{2} & \Lambda^{2}{}_{2} & \Lambda^{3}{}_{2} \\ -\Lambda^{0}{}_{3} & \Lambda^{1}{}_{3} & \Lambda^{2}{}_{3} & \Lambda^{3}{}_{3} \end{pmatrix} \\ &\equiv \left(\Lambda_{\nu}{}^{\mu}\right) = \begin{pmatrix} \Lambda_{0}{}^{0} & \Lambda_{0}{}^{1} & \Lambda_{0}{}^{2} & \Lambda_{0}{}^{3} \\ \Lambda_{1}{}^{0} & \Lambda_{1}{}^{1} & \Lambda_{1}{}^{2} & \Lambda_{1}{}^{3} \\ \Lambda_{2}{}^{0} & \Lambda_{2}{}^{1} & \Lambda_{2}{}^{2} & \Lambda_{2}{}^{3} \\ \Lambda_{3}{}^{0} & \Lambda_{3}{}^{1} & \Lambda_{3}{}^{2} & \Lambda_{3}{}^{3} \end{pmatrix} \\ &= \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{E.64}$$

As expected, $[\Lambda]^{-1}$ is obtained by simply changing the sign of β in Λ. We can write (E.62) in a matrix form as follows

$$\begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix}. \tag{E.65}$$

It is easy to verify that equations (E.59) and (E.62) lead, respectively, to

$$\Lambda^{\mu}{}_{\nu} = \frac{\partial x'^{\mu}}{\partial x^{\nu}}, \qquad \text{and} \qquad \Lambda_{\nu}{}^{\mu} = \left(\Lambda^{-1}\right)^{\mu}{}_{\nu} = \frac{\partial x^{\mu}}{\partial x'^{\nu}}; \tag{E.66}$$

these two relations, in turn, lead to

$$\Lambda^{\alpha}{}_{\mu}\Lambda_{\alpha}{}^{\nu} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}}\frac{\partial x^{\nu}}{\partial x'^{\alpha}} = \frac{\partial x^{\nu}}{\partial x^{\mu}} = \delta_{\mu\nu}. \tag{E.67}$$

Using the relation $\gamma^2 = 1/(1-\beta^2)$, we can verify that the determinants of the Lorentz transformation (E.61) and its inverse (E.64) are both equal to 1; namely,

$$\det([\Lambda]) = \det([\Lambda]^{-1}) = \gamma^2(1-\beta^2) = 1. \tag{E.68}$$

Invariance condition of the inner product between any two 4-vectors

Let us now look for the condition that will make the inner product between any two 4-vectors a Lorentz invariant; that is, we want to derive the condition that will make $A' \cdot B' = A \cdot B$, where A' and B' are the Lorentz transforms of A and B. Using Equations (E.36) and (E.59), we can write

$$A'_{\mu} = g_{\mu\nu}A'^{\nu} = g_{\mu\nu}\Lambda^{\nu}{}_{\alpha}A^{\alpha}, \qquad B'^{\mu} = \Lambda^{\mu}{}_{\beta}B^{\beta}, \tag{E.69}$$

which, in turn, lead to

$$A' \cdot B' = A'_{\mu}B'^{\mu} = \left(g_{\mu\nu}\Lambda^{\nu}{}_{\alpha}\Lambda^{\mu}{}_{\beta}\right)A^{\alpha}B^{\beta}. \tag{E.70}$$

On the other hand, invoking the relation $A_{\beta} = g_{\beta\alpha}A^{\alpha}$ (see (E.36)), we can write the inner product $A \cdot B$ as

$$A \cdot B = A_{\beta}B^{\beta} = g_{\beta\alpha}A^{\alpha}B^{\beta}. \tag{E.71}$$

A combination of (E.70) and (E.71) shows that in order for the relation $A' \cdot B' = A \cdot B$ (or for $A'_{\mu}B'^{\mu} = A_{\mu}B^{\mu}$)to hold for any two 4-vectors A and B, the coefficients of $A^{\alpha}B^{\beta}$ on the right hand sides of (E.70) and (E.71) must be equal: $g_{\mu\nu}\Lambda^{\mu}{}_{\beta}\Lambda^{\nu}{}_{\alpha} = g_{\beta\alpha}$; that is:

$$\boxed{A'_{\mu}B'^{\mu} = A_{\mu}B^{\mu} \quad \Longleftrightarrow \quad g_{\mu\nu}\Lambda^{\mu}{}_{\beta}\Lambda^{\nu}{}_{\alpha} = g_{\beta\alpha}.} \tag{E.72}$$

This relation provides the mathematical definition of all Lorentz transformations $\Lambda^{\mu}{}_{\nu}$. Any matrix $[\Lambda]$ obeying this equation is a valid Lorentz transformation. We can write (E.72) in a matrix form as follows:.

$$[\Lambda]^T[g][\Lambda] = [g], \tag{E.73}$$

where $[\Lambda]^T$ is the matrix transpose of $[\Lambda]$. This means that the metric remains unchanged under Lorentz transformations; the invariance of the metric under Lorentz transformations is a key property of the Minkowski spacetime (which is a *flat* spacetime). However, in a *curved* spacetime, such as in general relativity, the metric $g_{\mu\nu}(x)$ transforms into $g'_{\mu\nu}(x')$ such that $g'_{\mu\nu}(x') \neq g_{\mu\nu}(x)$.

Example E.1

Use the matrices of $[\Lambda]$, $[g]$, and $[\Lambda]^T$ to show that $[\Lambda]^T g[\Lambda] = [g]$.

Solution

First, equation (E.61) shows that $[\Lambda]$ is symmetrical and, hence, it is equal to its transpose:

$$[\Lambda]^T = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = [\Lambda]. \tag{E.74}$$

Second, using the matrices (E.35), (E.61) and (E.74), we can write

$$\begin{aligned}
[\Lambda]^T[g][\Lambda] &= \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ -\beta\gamma & -\gamma & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \gamma^2\left(1-\beta^2\right) & -\gamma^2\beta+\beta\gamma^2 & 0 & 0 \\ -\gamma^2\beta+\beta\gamma^2 & -\gamma^2\left(1-\beta^2\right) & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
&= [g],
\end{aligned} \tag{E.75}$$

where we have used the fact that $\gamma^2\left(1-\beta^2\right) = 1$.

E.7 Invariance under the Lorentz transformation

As stated in Subsection E.3, special relativity is based on the two premises that the laws of physics have the same form in all inertial (Lorentzian) frames[15] and that the speed of light is a constant for all observers, independent of their relative motion. In general, a quantity whose value is the same in all reference frames is called a *Lorentz invariant* or a Lorentz scalar[16]. For instance, the spacetime interval between any two events is independent of the inertial frame in which they are recorded.

[15]The covariance of the laws of physics in all inertial frames is an experimental fact. Since there is no universal reference frame, the physical laws must have *the same form* in all inertial frames; this is the essence of the *covariance* of the laws of physics.

[16]In addition to what was mentioned above, one needs to distinguish between "invariance" and "covariance." The *invariance* of a quantity means that it *does not change at all*; it remains the same in all reference frames. The *covariance* of an expression (*e.g.*, 4-vector or equation) means that it changes in a particular way under transformation from one system to another while *maintaining the same form*. So, while only scalar quantities can be invariant, 4-vectors (or 4-vector equations) can be, at most, covariant under Lorentz transformations.

As shown in Chapter 12, Lorentz transformations conserve the following quantities: (a) the length of a 4-vector, (b) the distance between any two events,(c) the causal nature of a 4-vector, and (d)the inner product between any two vectors. In what follows, we are going to prove that the inner product between any two 4-vectors is Lorentz invariant.

Invariance of the inner product between 4-vectors

We have seen in Chapter 7 that the inner product between any two vectors (in the 3-dimensional Euclidean space) is an invariant. Similarly, we can show that the inner product between any two 4-vectors is a Lorentz invariant:

$$A^0B^0 - A^iB^i = A'^0B'^0 - A'^iB'^i \qquad \longrightarrow \qquad A^\mu B_\mu = A'^\mu B'_\mu, \tag{E.76}$$

where A'^μ and B'_μ are the Lorentz transforms of A^μ and B_μ; notice that we have used the Einstein summation convention with respect to the indices i and μ with $i = 1,2,3$ and $\mu = 0,1,2,3$. It is worth mentioning that, as shown in Chapter 12 (see Example 12.2), the quantity $A^\mu B^\mu = A^0B^0 + A^iB^i$ is not invariant under Lorentz transformations; its numerical value will not be the same in different inertial frames.

Since the inner product invariance (E.76) is a direct algebraic consequence of the Lorentz transformation (E.2), it should be valid for all 4-vectors. We can list three noteworthy applications of (E.76) that will be of particular relevance to our presentation when we deal with the Klein-Gordon and the Dirac equations below:

- The First application of (E.76) is to prove the invariance of the magnitude square of the 4-displacement vector x^μ; namely, we have

$$(ct)^2 - \left(x^2 + y^2 + z^2\right) = (ct')^2 - \left(x'^2 + y'^2 + z'^2\right) \qquad \longrightarrow \qquad x^\mu x_\mu = x'^\mu x'_\mu. \tag{E.77}$$

- The second application of (E.76) is to the calculation of the inner product between two 4-momentum vectors p^μ, which is given by $p^\mu p_\mu = E^2/c^2 - \vec{p}^{\,2}$. It has the same value in all inertial frames. For instance, in the frame of reference in which a particle is at rest (i.e., $\vec{p} = 0$), the energy E reduces to the rest-mass energy, $E = mc^2$, where m is the rest mass[17] of the particle. Since the mass of a particle is the same in all frames (very much like the speed of light), the magnitude square of the 4-momentum is a Lorentz scalar which is equal to mc^2:

$$\boxed{p^\mu p_\mu = p^{0^2} - \vec{p}^{\,2} = \frac{E^2}{c^2} - \vec{p}^{\,2} = mc^2.} \tag{E.78}$$

- Finally, applying (E.76) to the inner product between the gradient operators ∂_μ and ∂^μ, we can easily verify that the d'Alembert operator is also a Lorentz invariant, since $\partial_\mu\partial^\mu$ is a scalar.

[17]The rest mass of a particle is the mass measured by an observer riding along with the particle (i.e., in the same frame).

Appendix F

Lagrangian Formulation of Classical Mechanics

This appendix is intended to give the reader the essential tools to follow the topic of *Lagrangians in Classical Field Theory* presented in Chapter 13. We are going to recall some of the most essential elements of classical mechanics. Classical mechanics can be formulated in three ways: the Newtonian, the Lagrangian (or analytical), and the Hamiltonian (or canonical) formulations. We will offer below brief reviews of the Lagrangian and the Hamiltonian formulations in Subsections F.1 and F.2, respectively. As for the Newtonian formulation, a topic that is too elementary to cover in this text, it suffices to mention that, the motion of a single particle of mass m and position $\vec{r}$ moving under the influence of a force, $\vec{F}$, that is derived from a potential $V(\vec{r})$, $\vec{F} = -\vec{\nabla}V(\vec{r})$ (i.e., a conservative force), is governed by Newton's second law:

$$m\frac{d^2\vec{r}}{dt^2} = -\vec{\nabla}V(\vec{r}). \tag{F.1}$$

As we will see below (see (F.13) and (F.37)), the Lagrangian and Hamiltonian formulations provide alternative ways for obtaining this equation.

F.1 Lagrangian Formulation of Classical Mechanics

Newton introduced his theory of classical mechanics in his 1687 masterpiece *Philosophiae Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy). Almost a century later, in 1788, Lagrange published his book "Mécanique Analytique" (*Analytical Mechanics*) in which he has reformulated Newtonian mechanics from a more fundamental approach using variational calculus which, in the end, led to a simple second-order differential equation. In what follows, we are going to highlight the most essential elements of the Lagrangian formalism. First, let us take a quick look at the generalized coordinates; they are known to greatly simplify the Lagrangian and Hamiltonian formalisms.

A system of N particles with m holonomic constraints

Consider, a system of N particles in three dimensional space, each with a position vector $\vec{r}_j(t)$, where $j = 1, 2, 3, \cdots, N$; each $\vec{r}_j(t)$ is a 3-vector, $\vec{r}_j(t) = (x_j, y_j, z_j)$. Hence, we have

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.
© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

a total of $3N$ coordinates: $\vec{r}_1(t) = (x_1, y_1, z_1)$, $\vec{r}_2(t) = (x_2, y_2, z_2)$, $\cdots$, $\vec{r}_N(t) = (x_N, y_N, z_N)$. According to Newtonian mechanics, the motion of each particle is described by the second law, $\vec{F}_j = m_j d^2\vec{r}_j/dt^2$. If there are no constraints on the system, we would need $3N$ coordinates to specify the system; this is known as the *configuration space*. However, when the N particles are subjected to constraints, we would need less than $3N$ coordinates to describe the system, since the constraints will limit its motion.

Constraints can be classified in two groups: *holonomic* and *nonholonomic*. Holonomic constraints only involve the coordinates of the configuration space and can be expressed in the form of an equation that connects the coordinates of the various particles of the system: $f(\vec{r}_1, \vec{r}_2, \vec{r}_3, \cdots, \vec{r}_N, t) = 0$. Nonholonomic constraints, however, cannot be expressed in a similar form. We are going to consider holonomic constraints only.

Let us consider the case where the $N-$particle system has m holonomic constraints given by

$$f_k(\vec{r}_1, \vec{r}_2, \vec{r}_3, \cdots, \vec{r}_N, t) = 0, \qquad k = 1, 2, 3, \cdots, m, \tag{F.2}$$

where $m < N$. As a result, some of the N position vectors will be dependent as we can use the constraints to eliminate m of the $3N$ coordinates. Hence, we will be left with $3N - m$ independent coordinates; the dimension of the configuration space would then become $3N-m$. The dimension of the configuration space is called the *number of degrees of freedom* of the system; they represent the number of independent coordinates. In other words, the number of the qs (which is equal to $n = 3N - m$) is the dimension of configuration space; it gives the number of degrees of freedom of the system.

As a simple illustration, let us take a rigid body which consists of, say, $N = 4$ particles; the distance between each two particles is constant: $|\vec{r}_i - \vec{r}_j| = constant$ where $i, j = 1, 2, 3, 4$. In this particular case, and avoiding double counting, we will have $m = 6$ holonomic constraints; hence, since each particles is specified by 3 coordinates, the number of degrees of freedom is: $n = 3$ (coordinates) $\times$ 4 (particles) -6 (constraints) $= 6$. So, this rigid body of 4 particles will have $n = 6$ degrees of freedom (i.e., 6 independent coordinates).

Generalized Coordinates

To describe the system of N particles with m holonomic constraints (i.e., with $n = 3N - m$ degrees of freedom), it more suitable to use a set of n independent coordinates $q_1, q_2, \cdots, q_n$ such that each position vector $\vec{r}_j(t)$ is expressed in terms of the n coordinates q_j (where $j = 1, 2, 3, \cdots, n$) and possibly time as follows:

$$\begin{aligned}
\vec{r}_1 &= \vec{r}_1(q_1, q_2, q_3, \cdots, q_n, t) \\
\vec{r}_2 &= \vec{r}_2(q_1, q_2, q_3, \cdots, q_n, t) \\
\vdots &= \vdots \\
\vec{r}_N &= \vec{r}_N(q_1, q_2, q_3, \cdots, q_n, t).
\end{aligned} \tag{F.3}$$

Conversely, we have

$$\begin{aligned}
q_1 &= q_1(\vec{r}_1, \vec{r}_2, \vec{r}_3, \cdots, \vec{r}_N, t) \\
q_2 &= q_2(\vec{r}_1, \vec{r}_2, \vec{r}_3, \cdots, \vec{r}_N, t) \\
\vdots &= \vdots \\
q_n &= q_n(\vec{r}_1, \vec{r}_2, \vec{r}_3, \cdots, \vec{r}_N, t).
\end{aligned} \tag{F.4}$$

The set $q_1, q_2, \cdots, q_n$ are called the *generalized coordinates*[1]. Moreover, the *generalized velocities* are written as $\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n$, where $\dot{q}_j = dq_j/dt$ with $j = 1, 2, 3, \cdots, n$. Sometimes, we will be using the shorthand notations q and $\dot{q}$ to designate: $q \equiv (q_1, q_2, \cdots, q_n)$ and $\dot{q} \equiv (\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n)$.

Euler-Lagrange Equations

The *Lagrangian* of a system of N non-relativistic particles (of masses $m_1, m_2, \cdots, m_N$) that are subjected to m holonomic constraints is defined as the difference of kinetic and potential energies of the system:

$$L(q, \dot{q}; t) = T - V \qquad \text{where} \qquad T = \frac{1}{2}\sum_{j=1}^{n} m_j \dot{q}_j^2, \quad V = V(q), \tag{F.5}$$

where $q \equiv (q_1, q_2, \cdots, q_n)$ and $\dot{q} \equiv (\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n)$. Here, $q_j(t)$ and $\dot{q}_j(t)$ are the dependent variables, while time is the independent variable. Clearly, the Lagrangian formalism is based on energies, not forces as in the case of Newtonian mechanics; as we will below, this makes the Lagrangian a scalar quantity and, hence, an invariant under transformations.

The main motivation in taking the Lagrangian in the form (F.5) is to recover Newton's second law from the Lagrangian formalism (see Equation (F.13)). So, since Newton's second law for each particle is a second-order differential equation, $\vec{F}_j = m_j d^2\vec{r}_j/dt^2$, the Lagrangian has to depend on two variables, q and $\dot{q}$; hence $L(q, \dot{q}; t)$ is a function of the coordinates, the velocities and the time, since q and $\dot{q}$ depend on time. Additionally, we are going to restrict our discussion to the case where the N–particle system is under the influence of *purely conservative forces*; i.e., forces that are derivable from the gradient of scalar potentials. When an object moves from one location to another under the influence of a conservative force, the change of the potential energy of the object does not depend on the path taken; it only depends on the initial and final points. As such, the potential energy corresponding to a *conservative force* is *velocity-independent* and does not depend on time either; that is why we took $V(q)$ to depend on the generalized coordinates $q = (q_1, q_2, \cdots, q_n)$ but not on the generalized velocities $\dot{q} = (\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n)$ nor on time (as a reminder, for non-conservative forces, the potential $V(q, t)$ depends explicitly on time). It is worth noting that the Lagrangian has the dimensions of energy. In sum, the form (F.5) fulfills all the needed requirements.

The *Hamilton's principle of least action* of classical mechanics asserts that the motion of the N–particle system from time t_1 to t_2 is such that the action (integral of the Lagrangian), S, has a stationary value

$$\delta S = \delta \int_{t_1}^{t_2} dt\, L(q, \dot{q}; t) = 0, \tag{F.6}$$

where $L(q, \dot{q}; t) \equiv L(q_1, q_2, q_3, \cdots, q_{3N-m}, \dot{q}_1, \dot{q}_2, \cdots, \dot{q}_{3N-m}, t)$. The resulting equations of motion for this system of N particles with m holonomic constraints are given by the *Euler-*

[1]The coordinates q_j are called *generalized* because they can be any set of coordinates that are not necessarily rectangular, like the *Curvilinear* coordinates that are used in a coordinate system for Euclidean space where the unit base vectors do not have to be orthogonal and where the coordinate axes may be *curved*. Polar coordinates, like the cylindrical and spherical coordinates, are examples of curvilinear coordinates, yet their base vectors are orthogonal.

Lagrange equations:

$$\boxed{\frac{\partial L}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = 0, \qquad j = 1, 2, 3, \cdots, 3N - m.} \tag{F.7}$$

One of the advantages of the Lagrangian formalism is the fact that it is so flexible and versatile that the Euler-Lagrange equations of motion hold their forms in any configuration space.

Deriving Newton's Second Law from the Lagrangian

For simplicity, let us illustrate the derivation on a single particle with no constraints. The *Lagrangian* for a non-relativistic, single particle of mass m in a 3-dimensional space can be easily inferred from (F.5) by simply taking $n = 3$, since a particle has 3 coordinates (and, hence, 3 degrees of freedom):

$$L(q, \dot{q}; t) = T - V \qquad \text{where} \qquad T = \frac{1}{2}m\sum_{i=1}^{3}\dot{q}_j^2, \quad V = V(q), \tag{F.8}$$

where $q = (q_1, q_2, q_3)$ and $\dot{q} = (\dot{q}_1, \dot{q}_2, \dot{q}_3)$. For instance, in the case of Cartesian coordinates, we have: $q_1 = x$, $q_2 = y$, $q_3 = z$ and $\dot{q}_1 = v_x$, $\dot{q}_2 = v_y$, $\dot{q}_3 = v_z$. The equations of motion for this single particle can be obtained from (F.7):

$$\frac{\partial L}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = 0, \qquad j = 1, 2, 3. \tag{F.9}$$

Using (F.8), we can easily show that

$$\frac{\partial L}{\partial q_j} = -\frac{\partial V}{\partial q_j}, \qquad \frac{\partial L}{\partial \dot{q}_j} \equiv \frac{\partial T}{\partial \dot{q}_j} = m\dot{q}_j \quad \Longrightarrow \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) \equiv \frac{d}{dt}\left(m\dot{q}_j\right) = m\frac{d^2q_j}{dt^2}. \tag{F.10}$$

Inserting these expressions into (F.9), we obtain

$$-\frac{\partial V}{\partial q_j} = m\frac{d^2q_j}{dt^2}, \qquad j = 1, 2, 3. \tag{F.11}$$

In the case of a conservative force, $\vec{F}$, such as gravity, we can express $\vec{F}$ as the gradient of a potential energy V; for instance, in Cartesian coordinates (i.e., $q_1 = x$, $q_2 = y$, $q_3 = z$), we have

$$\vec{F} = -\frac{\partial V}{\partial x}\hat{i} - \frac{\partial V}{\partial y}\hat{j} - \frac{\partial V}{\partial z}\hat{k} \equiv -\vec{\nabla}V. \tag{F.12}$$

Combining (F.11) and (F.12), we see that the Euler-equation (F.9) reduces to Newton's second law (F.1):

$$-\nabla V(\vec{r}) = m\frac{d^2\vec{r}}{dt^2} \quad \Longrightarrow \quad \vec{F} = m\vec{a} \qquad \text{where} \qquad \vec{a} = a_x\hat{i} + a_y\hat{j} + a_z\hat{k}; \tag{F.13}$$

$\vec{a}$ is the particle's acceleration. An important property of the Lagrangian formulation is that it can be used to obtain the equations of motion of an object in any coordinate system, not just the Cartesian coordinates, since the Euler-Lagrange equation (F.7) applies to any sets of coordinates.

In addition to the above derivation of Newton's second law from the Euler-Lagrange equation, one can also easily derive the Euler-Lagrange equation from Newton's second. This derivation is provided in standard classical mechanics textbooks[2].

Remarks

- Note that the form (F.5) of the Lagrangian is not unique: if we multiply $L(q, \dot{q}; t) = T - V$ by a constant and/or add a constant to it, the physics will not change, since these constants will cancel out in the Euler-Lagrange equation. For instance, the Lagrangians L and $L' = aL + b$, where a and b are constants, will satisfy the same equation of motion (F.7): since $\partial L'/\partial q_j = a\partial L/\partial q_j$ and $\partial L'/\partial \dot{q}_j = a\partial L/\partial \dot{q}_j$ and when we apply L' to (F.7), the constant a will cancel out.

- Similarly, one can also show the invariance of the equations of motion under a shift of the Lagrangian by a total time derivative of a function: $L \longrightarrow L + df(t)/dt$, where $f(t)$ is a time-dependent function. This is easy to see, since the variation of the action δS would generate a term, $\delta \int_{t_1}^{t_2} dt\,[df(t)/dt]$, that vanishes at the end points: $\delta \int_{t_1}^{t_2} dt\,[df(t)/dt] = \delta f(t_2) - \delta f(t_1) = 0$ because $f(t_1) = f(t_2) = 0$.

Advantages of the Lagrangian Formulations of Classical Mechanics

As outlined above, we have seen how to obtain Newton's second law from the Euler-Lagrange equation. It is important to note that, notwithstanding the fact that the Newtonian and the Lagrangian formulations of classical mechanics are equivalent, the Lagrangian approach has a number of advantages compared to Newton's laws:

- The Lagrangian formulation frees us from sticking or committing to any particular coordinate system. The generalized coordinates q can represent Cartesian, polar, cylindrical, or spherical coordinates, or any other sets of coordinates.

- The Lagrangian is a scalar function (the kinetic and potential energies are scalar quantities) and the Euler-Lagrange equations involve only scalar quantities, while the equations of Newton are vector equations with a number of components. As such, whereas the Euler-Lagrange equations are invariant under transformations to new coordinates systems, the various components of Newton's equations are not invariant; it is much easier to transform the single-component (scalar) Lagrangian than the multi-component (vector) equations of Newton.

- The Lagrangian formulation handles symmetries with ease and clarity: it provides an effective tool to connect symmetries with conservation laws. If the Lagrangian is invariant under a symmetry, the resulting equations of motion are also invariant under that particular symmetry. As illustrated in Table 13.1, it is easy to see the connection between the symmetries of the Lagrangian and the existence of conserved quantities;

[2]For instance, see: H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd Edition, Pearson (2001).

every symmetry of the Lagrangian (system) has a corresponding conservation law (Noether's theorem). For instance, the invariance of the Lagrangian under spatial translations leads at once to the conservation of linear momentum; Newton's laws do not say much about conserved quantities for symmetries.

- An other noteworthy advantage of the Lagrangian formulation is the ease with which it can be extended to describe systems that are not considered in Newtonian mechanics. For instance, it can be extended to relativistic systems and to field theories (as will be seen in the following subsection) and, hence, to other areas of physics such as electromagnetism. Finally, the Lagrangian formulation lends itself quite readily to quantization, for it leads to the Hamiltonian formulation which, in turn, allows for a smooth transition to quantum mechanics through the well known procedure of canonical quantization (see Subsection F.3).

F.2 Hamiltonian Formulation of Classical Mechanics

Hamilton introduced his formulation of the *least* or *stationary action* principle in two separate works[3] that he published in 1834 and 1835. This principle, which later became known as *Hamilton's principle*, provides the foundation for the Hamiltonian formulation of classical mechanics. Let us go over the most essential steps of this formulation.

The transition from the Lagrangian formulation of classical mechanics to the Hamiltonian formulation is carried out means of the *Legendre transformation* that allows us to shift from the dynamical variables $(q_j,\ \dot{q}_j,\ t)$ to $(q_j,\ p_j,\ t)$, where p_j is the *generalized momentum conjugate* to q_j:

$$\boxed{p_j = \frac{\partial L}{\partial \dot{q}_j}\,.} \tag{F.14}$$

This relation is similar to the second equation in (F.10). Inserting (F.14) into (F.7), we can rewrite the Euler-Lagrange equation in terms of p_j as follows

$$\boxed{\frac{dp_j}{dt} = \frac{\partial L}{\partial q_j} \quad \longrightarrow \quad \dot{p}_j = \frac{\partial L}{\partial q_j}.} \tag{F.15}$$

First, let us begin by taking the total differentiation of the Lagrangian, $L(q_j,\ \dot{q}_j;\ t)$:

$$dL = \sum_j \left(\frac{\partial L}{\partial q_j} dq_j + \frac{\partial L}{\partial \dot{q}_j} d\dot{q}_j \right) + \frac{\partial L}{\partial t} dt, \tag{F.16}$$

which leads to the following total derivative

$$\frac{dL}{dt} = \sum_j \left(\frac{\partial L}{\partial q_j} \frac{dq_j}{dt} + \frac{\partial L}{\partial \dot{q}_j} \frac{d\dot{q}_j}{dt} \right) + \frac{\partial L}{\partial t}. \tag{F.17}$$

[3] W.R. Hamilton, "On a General Method in Dynamics," *Philosophical Transactions of the Royal Society*, Part II (1834) pp. 247–308; W.R. Hamilton, "Second Essay on a General Method in Dynamics," *Philosophical Transactions of the Royal Society*, Part I (1835) pp. 95–144.

Next, let us consider the case where the Lagrangian describes a system that does not interact with anything outside, i.e., a *closed system*. In this case, *time is homogeneous* and, hence, the *Lagrangian does not depend explicitly on time*:

$$\frac{\partial L}{\partial t} = 0. \tag{F.18}$$

Hence, (F.17) becomes

$$\frac{dL}{dt} = \sum_j \left(\dot{q}_j \frac{\partial L}{\partial q_j} + \ddot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right), \tag{F.19}$$

where $\ddot{q}_j = d\dot{q}_j/dt$. On the other hand, the Euler-Lagrange equation (F.7) yields

$$\frac{\partial L}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j} \right), \tag{F.20}$$

which, when inserted in (F.19), leads to

$$\frac{dL}{dt} = \sum_j \left[\dot{q}_j \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j} \right) + \ddot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right] \quad \Longrightarrow \quad \frac{dL}{dt} = \frac{d}{dt}\left(\sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right), \tag{F.21}$$

or to

$$\frac{d}{dt}\left(L - \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) = 0 \quad \Longrightarrow \quad L - \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} = \text{constant}. \tag{F.22}$$

Let us denote this constant by $-H$; i.e., $L - \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} = -H$. Thus, we obtain

$$H = \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L. \tag{F.23}$$

This function is called the *Hamiltonian* of the system. It has the same dimensions as those of the Lagrangian; that is, the dimensions of energy. In what follows, we are going to show that the Hamiltonian represents the total energy of the system, provided several conditions are met.

As mentioned above, if the potential V does not depend explicitly on the velocities $\dot{q}_j$ or time, then $V = V(q)$ and, hence, $\partial V/\partial \dot{q}_j = 0$. Additionally, since $L = T - V$, we can write

$$\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial (T - V)}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j}. \tag{F.24}$$

On the other hand, if the kinetic energy is a homogeneous quadratic function of $\dot{q}_j$, we have $T \equiv T(\dot{q}_j) = \frac{1}{2} m \sum_j \dot{q}_j^2$ (see Equation (F.5)); thus, we can rewrite (F.24) as

$$\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j} = m\dot{q}_j. \tag{F.25}$$

Inserting this relation into (F.23) and using the fact that $m\sum_j \dot{q}_j^2$=2T, we obtain

$$H = 2T - L, \tag{F.26}$$

which, when combined with $L = T - V$, leads to

$$H = T + V \equiv E \qquad \Longrightarrow \qquad E = \text{constant}. \tag{F.27}$$

So, the total energy E is a constant of the motion and equal to the Hamiltonian provided that these two conditions are satisfied: (a) the potential energy does not depend on the generalized velocity (i.e., $\partial V/\partial \dot{q}_j = 0$) nor on time[4], and (b) the kinetic energy is a homogeneous quadratic function of $\dot{q}_j$. But if the potential energy depends on time, then $T + V$ is not a constant of the motion.

Lagrangian Symmetries and the Corresponding Conserved Quantities

As mentioned above, one can handle symmetries and conservation laws with great ease within the Lagrangian formalism. As illustrated in Table 13.1, let us look at two such symmetries. First, we have shown above that, if the Lagrangian does not depend on time explicitly as displayed in Equation (F.18), the total energy is constant as evidenced by (F.27). Second, if the Lagrangian is invariant under translations, we would expect the partial derivative of the Lagrangian with respect to the generalized coordinate q_j to vanish,

$$\frac{\partial L}{\partial q_j} = 0. \tag{F.28}$$

Inserting this relation into the Euler-Lagrange equation (F.7), we obtain

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = 0 \qquad \Longrightarrow \qquad \frac{\partial L}{\partial \dot{q}_j} = \text{constant}, \tag{F.29}$$

which, when combined with $L = T - V = \frac{1}{2}m\sum_{i=1}^{3}\dot{q}_j^2 - V$, leads to

$$\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial (T-V)}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j} = \frac{\partial}{\partial \dot{q}_j}\left(\frac{1}{2}m\sum_{i=1}^{3}\dot{q}_j^2\right) = m\dot{q}_j \equiv p_j. \tag{F.30}$$

Hence, combining (F.29) with (F.30), we see that the generalized linear momentum is conserved: p_j = constant.

Hamilton's Equations of Motion

Let us now derive the Hamiltonian Equations of Motion. Instead of the Lagrangian $L(q, \dot{q}; t)$, we will work with the Hamiltonian, $H(q, p; t)$; the conversion of $L(q, \dot{q}; t)$ into $H(q, p; t)$ is carried out by means of the Legendre transformation. Using (F.14) and after inserting $\frac{\partial L}{\partial \dot{q}_j} = p_j$ into (F.23), we can write the Hamiltonian as

$$\boxed{H(q, p; t) = \sum_{j=1}^{n}\dot{q}_j p_j - L(q, \dot{q}; t),} \tag{F.31}$$

[4]That is, the relations (F.3) and (F.4) must not depend explicitly on time; in other words, $\vec{r}_j$ and q_j must be of the form: $\vec{r}_j = \vec{r}_j(q_1, q_2, q_3, \cdots, q_n)$ and $q_j = q_j(\vec{r}_1, \vec{r}_2, \vec{r}_3, \cdots, \vec{r}_N)$.

where, as seen above, $n = 3N-m$, $q \equiv (q_1, q_2, \cdots, q_n)$ and $p \equiv (p_1, p_2, \cdots, p_n)$. Let us calculate the total differentiation of H in two ways. First, by writing the differentiation of $H(q, p; t)$ as

$$dH(q, p; t) = \sum_j \left(\frac{\partial H}{\partial q_j} dq_j + \frac{\partial H}{\partial p_j} dp_j \right) + \frac{\partial H}{\partial t} dt, \tag{F.32}$$

and, second, by writing the differentiation of $H(q, p; t) = \sum_j \dot{q}_j p_j - L(q, \dot{q}; t)$ as

$$\begin{aligned} dH &= d\left(\sum_j \dot{q}_j p_j \right) - dL(q, \dot{q}; t) \\ &= \sum_j \left(\dot{q}_j dp_j + p_j d\dot{q}_j \right) - \sum_j \left(\frac{\partial L}{\partial q_j} dq_j + \frac{\partial L}{\partial \dot{q}_j} d\dot{q}_j \right) - \frac{\partial L}{\partial t} dt. \end{aligned} \tag{F.33}$$

Since $\frac{\partial L}{\partial q_j} = \dot{p}_j$ and $\frac{\partial L}{\partial \dot{q}_j} = p_j$, we can rewrite the last line of (F.33) as

$$\begin{aligned} dH &= \sum_j \left(\dot{q}_j dp_j + p_j d\dot{q}_j \right) - \sum_j \left(\dot{p}_j dq_j + p_j d\dot{q}_j \right) - \frac{\partial L}{\partial t} dt \\ &\equiv \sum_j \left(\dot{q}_j dp_j - \dot{p}_j dq_j \right) - \frac{\partial L}{\partial t} dt, \end{aligned} \tag{F.34}$$

where we have used the fact that the second and fourth terms in the first equation cancel out. Comparing (F.32) to the last line of (F.34), we obtain

$$\boxed{\dot{q}_j = \frac{\partial H}{\partial p_j}, \qquad \dot{p}_j = -\frac{\partial H}{\partial q_j},} \tag{F.35}$$

and

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{F.36}$$

Equations (F.35) are known as the *Hamilton's equations* or as the *canonical equations of motion*. Note that these equations are equivalent to the Euler-Lagrange equations (F.7); in fact,the Euler-Lagrange equations imply Hamilton's equations of motion. The Hamilton's equations, which consist of $2N$ first-order equations of motion, exhibit a clear symmetry between q and p.

In his reformulation of classical mechanics, Hamilton has essentially reduced the second-order Euler-Lagrange equation (F.7) to two first-order differential equations (F.35); in the process, the Lagrangian $L(q, \dot{q})$ is replaced with the Hamiltonian $H(q, p)$ as shown in (F.31). We should also mention that the Hamiltonian formulation of classical mechanics has a noteworthy disadvantage: it breaks the covariant nature of the Lagrangian formalism and, hence, encounters problems with relativistic theories because the Hamiltonian is not a Lorentz scalar[5]. The

[5]The Hamiltonian is the zero or time component of the 4-vector momentum, $H = p^0$. One needs to keep in mind that both the zero-component, p^0, and the other three components, p^j ($j = 2, 3, 4$), of the 4-momentum vector are the generators of the Lorentz group; they change under Lorentz transformations. Hence, the Hamiltonian is not Lorentz invariant.

Lagrangian, however, being a scalar quantity (it is invariant under Lorentz transformations), is readily amenable to relativistic theories.

Summary of Hamilton's Formulation of Classical Mechanics
To obtain Hamilton's canonical equations of motion for a system of N particles, one needs to follow these steps:

- Select the generalized coordinates, $q \equiv (q_1, q_2, \cdots, q_N)$, and identify the m constraints of the system, if any.
- Construct the Lagrangian: $L(q, \dot{q}; t) = T - V \equiv \frac{1}{2}\sum_{j=1}^{n} m_j \dot{q}_j^2 - V(q)$, where $n = 3N - m$.
- Obtain the generalized momenta, $p_j = \partial L/\partial \dot{q}_j$.
- Construct the Hamiltonian: $H(q, p; t) = \sum_j \dot{q}_j p_j - L(q, \dot{q}; t)$, where $p \equiv (p_1, p_2, \cdots, p_N)$.
- Obtain Hamilton's canonical equations: $\dot{q}_j = \partial H/\partial p_j, \quad \dot{p}_j = -\partial H/\partial q_j$.

Obtaining Newton's Second Law from Hamilton's Canonical Equations
We should note that, if we combine the two equations (F.35), we would recover Newton's second law (F.1). For this, we simply need to take the time derivative of the first equation in (F.35) and then make use of the second one. That is, since $\partial H/\partial p_j = p_j/m_j$, the time derivative of the first equation yields $\ddot{q}_j = \dot{p}_j/m_j$; then using the second equation along with the fact that $\partial H/\partial q_j \equiv \vec{\nabla}_j H$ and using Cartesian coordinates (for simplicity), we have:

$$\ddot{q}_j = \frac{1}{m}\dot{p}_j \implies m_j \ddot{q}_j = \dot{p}_j = -\frac{\partial H}{\partial q_j} \implies m_j \ddot{\vec{r}}_j = -\vec{\nabla}_j H \implies m_j \frac{d^2\vec{r}_j}{dt^2} = -\vec{\nabla}_j V(\vec{r}), \tag{F.37}$$

where we have used the fact that $\vec{\nabla}_j H = -\vec{\nabla}_j L$ and that $\vec{\nabla}_j L = -\vec{\nabla}_j V(\vec{r})$, since $L = T - V$ and that the kinetic energy, T, does not depend on $\vec{r}$ (only V depends on $\vec{r}$); the quantity $\vec{\nabla}_j V(\vec{r})$ stands for $\vec{\nabla}_j V(\vec{r}) \equiv \vec{\nabla}_j V(\vec{r}_1, \cdots, \vec{r}_j, \cdots, \vec{r}_N) = \frac{\partial V}{\partial x_j}\hat{i} + \frac{\partial V}{\partial y_j}\hat{j} + \frac{\partial V}{\partial z_j}\hat{k}$.

Energy Conservation
Let us now discuss the conservation of energy. First, let us consider a *dynamical observable* A that depends on the generalized coordinates, q_j , momenta, p_j, and possibly time, t; i.e., $A = A(q_j, p_j, t)$. The total time derivative of A is given by

$$\frac{dA}{dt} = \sum_j \left(\frac{\partial A}{\partial q_j}\frac{\partial q_j}{\partial t} + \frac{\partial A}{\partial p_j}\frac{\partial q_j}{\partial t}\right) + \frac{\partial A}{\partial t} = \sum_j \left(\frac{\partial A}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial A}{\partial p_j}\frac{\partial H}{\partial q_j}\right) + \frac{\partial A}{\partial t}, \tag{F.38}$$

which, using the Hamilton equations $\frac{\partial q_j}{\partial t} = \frac{\partial H}{\partial p_j}$ and $\frac{\partial p_j}{\partial t} = -\frac{\partial H}{\partial q_j}$ (see (F.35)), can be rewritten as follows

$$\frac{dA}{dt} = \sum_j \left(\frac{\partial A}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial A}{\partial p_j}\frac{\partial H}{\partial q_j}\right) + \frac{\partial A}{\partial t}, \tag{F.39}$$

or as

$$\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t}, \tag{F.40}$$

where, as introduced in Chapter 3 (see Subsection 3.8.1), the quantity $\{A, H\}$ is the *Poisson bracket* between A and H:

$$\{A, H\} = \sum_j \left(\frac{\partial A}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial A}{\partial p_j} \frac{\partial H}{\partial q_j} \right). \tag{F.41}$$

If the observable A depends on t, it is said to have an *explicit time-dependence*; i.e., $A = A(q_j, p_j, t)$. However, if A does not depend explicitly on time, but depends on t only through $q_j(t)$ and $p_j(t)$, the observable is said to have an *implicit time-dependence*; i.e., $A = A(q_j, p_j)$. In this context, the first terms $\{A, H\}$ in (F.40) clearly represents the implicit time-dependence of A, whereas the last term $\partial A/\partial t$ represents the explicit time-dependence. Thus, using Equation (F.40) and if A does not depend explicitly on time, we have $\partial A/\partial t = 0$ and, hence, $dA/dt = \{A, H\}$:

$$\text{If} \quad \frac{\partial A}{\partial t} = 0 \quad \Longrightarrow \quad \frac{dA}{dt} = \{A, H\}. \tag{F.42}$$

Now, if $dA/dt = 0$ or $\{A, H\} = 0$, the dynamical observable A is said to be a *constant of the motion*. That is, an observable A is a constant of the motion if its Poisson bracket with the Hamiltonian is zero.

We are now ready to discuss the condition that the conservation of energy. If we set $A = H$ in Equation (F.41), the Poisson bracket of H with itself vanishes: we have $\{A, H\} = \{H, H\} \equiv 0$. Hence, Equation (F.40) becomes

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}. \tag{F.43}$$

A combination of (F.36) and (F.43) leads at once to

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{F.44}$$

This shows that, if the Lagrangian does not depend explicitly on time[6], the Hamiltonian also does not depend explicitly on time; in this case, the Hamiltonian is a constant of the motion. So, if the Hamiltonian does not depend explicitly on time, we have $\partial H/\partial t = 0$ and, hence, the Hamiltonian is a conserved quantity, $dH/dt = 0$. What about energy conservation? Evidently, the energy will also be conserved provided that $H = E$. As seen in (F.27), the Hamiltonian equals the total energy only for those systems where the potential energy does not depend on the velocity, $\dot{q}$, and the kinetic energy is a homogenous quadratic function of $\dot{q}$.

F.3 Canonical Quantization: from Classical to Quantum Mechanics

Canonical quantization is a prescription that allows the transition from classical to quantum mechanics; namely, the transition from the classical mechanical Hamiltonian, $H(q, p)$, to the

[6] As a reminder, for conservative forces, the Lagrangian does not depend explicitly on time, $L \equiv L(q, \dot{q})$; however, for non-conservative forces, the Lagrangian depends explicitly on time, $L \equiv L(q, \dot{q}, t)$, due, for instance, to a time dependent potential $V(q, t)$.

quantum mechanical Hamiltonian, $\hat{H}$. In Chapter 3 (see Subsection 3.8.1), we saw how to go from quantum to classical mechanics; here, we are going to do the opposite, going from classical to quantum mechanics.

Using the *Poisson bracket* between two *dynamical observables* $A(q,\ p)$ and $B(q,\ p)$ where $q = (q_1,\ q_2,\ \cdots,\ q_n)$ and $p = (p_1,\ p_2,\ \cdots,\ p_n)$

$$\{A, B\} = \sum_{j=1}^{n} \left(\frac{\partial A}{\partial q_j} \frac{\partial B}{\partial p_j} - \frac{\partial A}{\partial p_j} \frac{\partial B}{\partial q_j} \right); \tag{F.45}$$

since the canonical variables q_j and p_j do not depend on each other (i.e., $\partial q_j/\partial p_k = 0$ and $\partial p_j/\partial q_k = 0$), this relation yields at once to the following standard Poisson bracket relations:

$$\{q_j,\ q_k\} = \{p_j,\ p_k\} = 0, \qquad \{q_j,\ p_k\} = \delta_{jk}. \tag{F.46}$$

If the generalized coordinates q_j and momenta p_j do not depend explicitly on time, and using Equation (F.40), we can rewrite Hamilton's equations (F.35) as follows:

$$\frac{dq_j}{dt} = \{q_j,\ H\}, \qquad \frac{dp_j}{dt} = \{p_j,\ H\}. \tag{F.47}$$

These equations represent the Poisson bracket formulation of Hamilton's equations of motion for classical mechanics. In fact, combining equations (F.42) and (F.47), we see that the time evolution of a dynamical variable A is given by the following Poisson bracket:

$$\frac{dA}{dt} = \{A, H\}. \tag{F.48}$$

The standard procedure of canonical quantization is carried out by means of an "algorithm" involving few steps:

- Replace the pair of classical canonical variables $(q,\ p)$ with a pair of *operators* $(\hat{q},\ \hat{p})$.
- Replace the Poisson brackets (F.46) with *canonical commutation relations* between operators

  $$[\hat{q}_j,\ \hat{q}_k] = [\hat{p}_j,\ \hat{p}_k] = 0, \qquad [\hat{q}_j,\ \hat{p}_k] = i\hbar\delta_{jk}, \tag{F.49}$$

  where $[\hat{A},\ \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ is the usual commutator between two operators $\hat{A}$ and $\hat{B}$.
- The dynamics of the system is regulated by operator equations corresponding to Hamilton's equations (F.47):

  $$\frac{d\hat{q}_j}{dt} = \frac{1}{i\hbar}[\hat{q}_j,\ \hat{H}], \qquad \frac{d\hat{p}_j}{dt} = \frac{1}{i\hbar}[\hat{p}_j,\ \hat{H}]. \tag{F.50}$$

  These are the equations of motion in the Heisenberg picture; as seen in Chapter 10 (see Subsection 10.2.2.1), the Heisenberg equation of motion for a dynamical variable $\hat{A}$ is given by

  $$\frac{d\hat{A}}{dt} = \frac{1}{i\hbar}[\hat{A},\ \hat{H}]. \tag{F.51}$$

  For instance, for the 3-dimensional position and momentum operators, we have

  $$\frac{d\hat{\vec{r}}_j}{dt} = \frac{1}{i\hbar}[\hat{\vec{r}}_j,\ \hat{H}], \qquad \frac{d\hat{\vec{p}}_j}{dt} = \frac{1}{i\hbar}[\hat{\vec{p}}_j,\ \hat{H}]. \tag{F.52}$$

To be more general, the transition from classical to quantum mechanics is carried out by implementing the following two steps:

1. Replace the classical dynamical variables with their corresponding operators of a Hilbert space,
$$A,\ B,\ C,\ \cdots \quad \longrightarrow \quad \hat{A},\ \hat{B},\ \hat{C},\cdots . \tag{F.53}$$

2. Replace the Poisson bracket between any pair of classical variables $\{A,\ B\}$ with a commutator $[\hat{A},\ \hat{B}]$ between the corresponding pair of operators and dividing it by $i\hbar$:
$$\{A, B\}_{classical} \quad \longrightarrow \quad \frac{1}{i\hbar}[\hat{A}, \hat{B}]. \tag{F.54}$$
One can easily obtain this relation by comparing (F.46) to (F.49), (F.47) to (F.50), and (F.48) to (F.51).

From the formalism outlined above, we see that there is a great deal of similarity between the *classical* Hamilton equations of motion in the Poisson bracket form (F.47) and the quantum mechanical equations of motion in the Heisenberg picture (F.50). This is precisely what has made the canonical quantization a practical and useful prescription to transition from classical to quantum mechanics. However, none can claim that this is a derivation of quantum mechanics from classical mechanics; this prescription only provides a quick, heuristic procedure to establish a connection between the two theories. Additionally, it is important to keep in mind that canonical quantization does not work for all quantum mechanical systems. For instance, it cannot handle the spin degree of freedom for it is a purely quantum mechanical effect. We know that the quantum mechanical Hamiltonian of a spin-1/2 particle has no classical analogue and, hence, cannot be obtained by means of a canonical quantization of some classical Hamiltonian. In cases like this one, we have to invoke the full fledged theory of quantum mechanics that doe not start from classical mechanics.

Now that we have seen how to apply the Lagrangian formalism to describe the motion of a discrete set of N particles, we are going to extend these (Lagrangian) concepts to the description of *fields*. So, instead of dealing with single-particle equations of motion, we are going to obtain *field equations*, for they are more appropriate to deal with relativistic quantum mechanics. In Chapter 13 we deal with this task; namely, an introduction to classical field theory.

Index

Abelian group, 92, 441
Absorption of a photon, 633
Action of a system, 900
Action quantization, 48
Addition of
    three angular momenta, 464–465
    two angular momenta, 448–464
Addition of angular momenta, 448–471
Addition of isospins, 470
Adiabatic approximation, 635
Adiabatic theorem, 635
Adjoint of an operator, 105
Airy functions, 304, 571
Algebraic method for
    harmonic oscillator, 266
Allowed transitions, 647
Alpha decay, 581
Angular momentum
    addition of three, 464–465
    addition of two, 448–464
    commutation relations, 314, 315
    conservation of, 210, 313, 375, 440
    eigenfunctions, 335–343
    general formalism, 315–320
    generator of infinitesimal rotations, 438
    geometrical representation, 323–324
    matrix representation, 320–323
    orbital, 313–315
    raising and lowering operators, 316
    spin, 325–333
Anisotropic harmonic oscillator, 372
Annihilation operator, 267–271
Anomalous Zeeman effect, 558
Anti-Hermitian operator, 106
Anticommutator, 108
Antilinear Operators, 114
Antineutrino, 787
Antiparticle, 25, 758
Antiprotons, 804
Antisymmetric state, 506, 511, 512
Antiunitary Operators, 114
Associated Laguerre polynomials, 392
Associated Legendre functions, 337
Atomic shell structure, 518
Axial vector, 471
Axial vectors, 805

Balmer series, 46
Barrier penetration, 254–258
Basis
    complete, 123, 124
    continuous, 141–148
    discrete, 123–140
    of a vector space, 94
    orthonormal, 94, 124, 141
Bessel equation, 378
Bessel functions, spherical, 378
Bianchi identity, 908, 929
Blackbody radiation, 13–19
Bohr atomic model, 41
Bohr energy, 43
Bohr magneton, 328, 401
Bohr quantization rule, 41
Bohr radius, 42
Bohr's postulates, 41
Bohr–Sommerfeld quantization rule, 573
Boltzmann constant, 16
Born approximation, first, 682
Born approximation, validity, 683
Born series, 680
Bose–Einstein condensation, 516
Bose–Einstein statistics, 511
Boson condensation, 516
Bosons, 510
Bound states, 242–243, 258–266

Quantum Mechanics: Concepts and Applications, Third Edition. Nouredine Zettili.

© 2022 John Wiley & Sons Ltd. Published 2022 by John Wiley & Sons Ltd.

Boundary conditions, 249
Box potential, 258
Bra-ket
  matrix representation of, 123
  notation, 97–103
  properties, 98

Canonical commutation relations, 147
Canonical equations of motion, 987
Center of mass motion, 302, 386
Central potential, 374–377
Centrifugal potential, 376
Characteristic equation, 138
Charge conjugation, 800
  of the pion, 804
  of the proton, 804
Chiral representation, 747
Chiral spinors, 785
Chirality, 786, 789
Chirality projection operators, 790
Classical limit, 214–215
Classical turning points, 242, 257, 568
Classically allowed region, 568
Classically forbidden region, 568
Clebsch–Gordan coefficients, 452
  orthonormalization relation, 452
  recursion relations, 455
Clebsch–Gordan series, 465–467
Clifford algebra, 737
  four dimensional, 741
  three-dimensional, 332, 737
Closed shell, 518
Closure relation, 124
Cofactor of a matrix, 128
Commutator
  algebra, 108–110
  anticommutator, 108
  definition, 108
  Jacobi identity, 109
  properties of a, 109
Compatible observables, 198
Complementarity principle, 36
Composite particles, 512
Compton effect, 23–25
Compton wavelength, 24
Configuration, 520
Connection formulas, 570–573
Conservation laws, 207–211
Conservation of
  energy, 209–210
  linear momentum, 209–210
  parity, 211
  probability, 204–205
Constant of the motion, 209–211, 440
Constrained variational principle, 561
Continuity of $\psi(x)$, 249
Continuity of $d\psi(x)/dx$, 249
Continuous spectrum, 243
Contravariant and Covariant notations, 712, 968
Contravariant and Covariant vectors, 712, 968
Correspondence principle, 215
Coulomb gauge, 400, 639, 640
Coulomb potential, 386, 684, 694, 701
Coulomb scattering, 684
Covariance of the
  Dirac equation, 749
  Klein–Gordon equation, 719
CPT theorem, 815
Creation operator, 267–271
Cross section
  Coulomb, 684
  differential, 671, 678, 682
  partial wave, 688
  total, 671, 688, 690
  total elastic, 689
  total inelastic, 690
Current density, 205, 218, 248, 755

d'Alembert operator, 718, 973
Darwin term, 553
Davisson–Germer experiment, 28
de Broglie relation, 28
de Broglie wavelength, 29
Degeneracy
  definition, 138
  exchange, 510–511
  for central potentials, 376
  for Coulomb potential, 396
  for cubical potential, 371
  for free particle, 369, 379
  for isotropic oscillator, 373, 384
  for symmetric potentials, 244
  of hydrogen levels, 396, 397

partial lifting of, 400
Degenerate perturbation theory, 544–547
Delta function, 141, 937
Delta potential, 285, 288
Density operator, 205
Detailed balancing, 634
Differential cross section, 671, 678, 682
Dipole moment, 478
Dipole selection rules, 647
Dipole–dipole interaction, 607
Dirac
hole theory, 764
pair annihilation, 764
pair creation, 764
Dirac algebra, 737
Dirac basis, 747
Dirac delta function, 141, 937
Dirac Equation, 734–817
continuity equation, 753
covariant form, 739–748
free-particle solutions, 756, 777
in an electromagnetic field, 793
Lagrangian derivation, 915
negative energies, 763
non-relativistic limit, 794–799
Dirac Matrices, 739
Dirac picture, 625
Dirac representation , 747
Discrete spectrum, 242–243
Distorted plane wave, 687
Double-slit experiment, 31–38
Dual vector space, 93, 97
Dynamical observables, 211
Dyson series, 628

Effective potential, 376
Ehrenfest theorem, 213–214
Eigenstate, 118
Eigenvalue, 138
Eigenvalue problem, 137–140
Eigenvalues and eigenvectors of an operator, 118–120
Eigenvectors, 138
Einstein summation convention, 709, 953, 971, 972
Einstein, photoelectic effect, 20
Elastic scattering
partial wave analysis, 685–689
total cross section, 689
Electric dipole approximation, 646
Electric dipole moment, 546, 647
Electric dipole transition, 646
Electromagnetic field quantization, 641–644
Electron cloud model, 409
Electronic configuration, 520
Energy conservation, 210
Equation
Klein–Gordon, 309
Equation of motion
density operator, 206
dynamical variable, 212
for expectation value, 206
for operators, 626
Hamilton–Jacobi, 214
Heisenberg, 625
interaction picture, 626
Euclidean metric, 710, 954
Euler angles, 442
Euler–Lagrange equation, 902
Euler-Lagrange equation, 982
Event in the Minkowski spacetime, 710, 964
Exchange degeneracy, 510–511
Exchange operator, 505
Exclusion principle, 516, 517
Expectation value, 196
time evolution, 206
Exponential decay, 254, 569

Fenyman Slash Notation, 743
Fermi golden rule, 631
Fermi–Dirac statistics, 511
Fermions, 510
Feynman-Stückelberg Interpretation, 766
Field strength tensor, 906
Fine structure constant, 42, 550
Fine structure of hydrogen, 552
flat spacetime metric, 958, 970
Flux, 205, 245, 248
Forbidden transitions, 647
Forward scattering amplitude, 688
Four vector, 965
four-current, 723
Fourier transforms, 51, 679
Free particle motion

one dimensional, 244–246
three dimensional, 369–370, 378–380

Galilean transformation, 962
Gauge field theory, 920
Gauge transformation, 904
Gauss's theorem, 561
Generalized coordinates, 211, 980, 981
Generator of
finite translations, 208
finite translations, 209
infinitesimal translations, 207, 208
infinitesimal rotations, 438
infinitesimal transformation, 207
Golden rule, 631
Goudsmit, Samuel, 326
Green's function, 679
Group velocity, 56–58
Gyromagnetic factor, 328
gyromagnetic factor, 796
Gyromagnetic ratio, 557
gyromagnetic ratio, 327, 796

Hamilton equations, 212
Hamilton's equations of motion, 987
Hamilton's principle, 902, 981
Hamilton–Jacobi equation, 214
Hankel functions, spherical, 381
Harmonic oscillator, 266–277
anisotropic, 372
energy eigenstates, 271–272
energy eigenvalues, 269–271
isotropic, 372
matrix representation, 275–276
wave function, 273
zero-point energy, 271
Heaviside function, 594
Heisenberg
equation of motion, 625
picture, 624–625
uncertainty principle, 38
uncertainty relations, 38, 111
Helicity, 782
Left-handed helicity, 783
Right-handed helicity, 783
Helium atom, energy levels, 529, 611, 612
Hermite polynomials, 267
Hermitian adjoint of an operator, 105–107
Hermitian operator, 106
Higgs boson, 725
Higgs boson particle, 708
Holonomic constraints, 980
Hund's rules, 522
Hydrogen atom, 385–399
anomalous Zeeman effect, 555–558
Bohr model, 41–48
degeneracy of energy levels, 396
energy levels, 389–391
fine structure, 552–553
normal Zeeman effect, 401–403
polarizability, 543
probabilities and averages, 397
radial functions, 391–394
radius quantization, 399
relativistic corrections, 551–552
spin–orbit coupling, 548–551
Stark effect, 543, 546
wave function, 394
Hyperfine structure, 553

Identical particles
indistinguishability of, 508–510
systems of, 508–516
Index Notation, 953
Inelastic scattering
partial wave analysis, 689–690
total cross section, 690
Inertial frame, 711, 962
Infinitesimal
rotations, 438
spatial translations, 208
time translations, 207
unitary transformations, 207
Interaction picture, 625–627
Interchange symmetry, 505
Intrinsic parity, 808
Invariance principle, 209
Invariance under
spatial rotation, 440
spatial translation, 210
time translation, 210
Inverse of a matrix, 127
Inverse of an operator, 113
Irreducible tensors, 474

Isospin, 468–471

Ket vector, 97, 98
Klein Gordon Equation, 715–734
  continuity equation, 723
  covariant continuity equation, 723
  covariant form, 718
  Lorentz Invariance, 719
  negative energies, 721
  negative probability densities, 722
  Non relativistic limit, 719
Klein Paradox, 730
Klein–Gordon equation, 309
  conceptual problems, 721
Klein-Gordon Equation
  Lagrangian derivation, 911
Kronecker delta, 124, 710, 953

Ladder method for
  angular momentum, 316–320
  harmonic oscillator, 266–271
Ladder operators, 269
Lagrange multipliers, 561
Lagrangian
  classical, 981, 982
  density, 900
Lagrangian density
  Dirac, 915
Laguerre polynomials, 392
Landé factor, 328, 557
Laplacian operator, 104
Larmor frequency, 403
Laser, 641
Least action principle, 902, 981
Legendre
  associated functions, 337
  differential equation, 337
  polynomials, 337
  polynomials, completeness, 337
Legendre transformation, 984
Levi–Civita tensor, 332
Levi-Civita tensor, 907, 927, 929, 956
Lifetime, 649
Light cone, 967
Lightlike, 713, 967
Linear operators, 105
Linear Stark effect, 543
Linear vector space, 91
Linearly independent vectors, 93
Lorentz boost, 975
Lorentz factor, 711, 962
Lorentz force, 923
Lorentz orthonormality condition, 713
Lorentz Transformation, 711
Lorentz transformation, 711, 962
Lowering operator, 269
Lyman series, 46

Magnetic dipole moment, 555
Majorana representation, 748
Majorana spinors, 748
Many-particle systems, 503–508
Maser, 641
Matrix
  cofactor of a, 128
  inverse of a, 127
  properties of a, 132
  skew-symmetric, 127
  symmetric, 127
  trace of a, 129
  transpose of a, 127
  unitary, 128
Matrix mechanics, 151
Matter waves, 30–31
Maxwell equations, 903
  homogeneous, 903
  inhomogeneous, 903
Mean lifetime, 649
Measurement, 195–201
Minimal coupling, 400, 405, 731, 925
Minimal substitution, 405, 731, 925
Minkowski metric, 713, 958, 970
Minkowski space, 964
Minkowski tensor, 958, 970
Mixed spectrum, 243
Momentum conservation, 210
Momentum representation, 143–144

negative energy, 721
Neumann algebras, 12
Neumann functions, spherical, 378
Neutrino, 786, 787
Noble gases, 523
Nodes

of a wave function, 243, 244, 260
of hydrogen radial functions, 394
of variational method wave function, 560
of WKB wave function, 574
Noether theorem, 899
Noether's theorem, 800, 899
Nonholonomic constraints, 980
Normalization
of associated Laguerre functions, 394
of associated Legendre functions, 339
of Gaussian wave packet, 54
of radial functions, 393
of spherical harmonics, 338
of WKB wave function, 574
Nuclear magneton, 797
Numerov algorithm, 278

Observables, 193–195
Occupation number, 268
Occupation number, 643
Old quantum theory, 8
Operator
Time reversal, 117
Operator density, 205
Operator theory, 12
Operators, 104–123
angular momentum, 315
Antilinear, 114
Antiunitary, 114
charge, 469
complete set of commuting, 198–200
definitions, 104
eigenvalues of, 118
even and odd, 149
exchange, 505
functions of, 112
Hermitian adjoint, 105–107
Inverse of an, 113
Laplacian, 104
linear, 105
matrix representation of, 126–131
parity, 148
permutation, 505
products of, 104
projection, 107
rotation, 440
scalar, 471
skew-Hermitian, 106
tensor, 473–476
trace of, 129
uncertainty of, 110
unitary, 113
vector, 472–473
Optical theorem, 689
Orbital
angular momentum, 313–315
magnetic dipole moment, 327, 401
orbital, 410
Orbitals, 397, 507, 518
Orthonormality condition, 124, 141

Pair annihilation, 764
Pair creation, 764
Pair production, 25–27
Parity, 786, 805
Parity conservation, 211
Parity operator, 148–150, 210
Parseval's theorem, 145
Partial wave analysis, 685
Paschen–Back effect, 556
Paschen–Back shift, 556
Pauli equation, 405, 406, 796
Pauli Exclusion Principle, 7
Pauli exclusion principle, 516–517
Pauli Hamiltonian, 334
Pauli matrices, 331
Pauli spinor, 405
Pauli Spinors, 331
Periodic table, 517–523
Perturbation theory
time-dependent, 627–634
time-independent, 538–558
degenerate, 544–547
nondegenerate, 538–544
Phase shift, 687
Phase velocity, 56–58
Photoelectric effect, 19–22
Picture
Heisenberg, 624–625
interaction, 625–627
Schrödinger, 624
Pictures of quantum mechanics, 623–627
Pion, 469, 487, 510, 512, 702, 804

Pion charge conjugation, 804
Planck's constant, 22
Planck's constant, 5
Planck's distribution, 17
Planck's postulate, 16
Poisson bracket, 989
Poisson brackets, 211–212
Polar vectors, 805
Polarizability, hydrogen atom, 543
Polynomials
  associated Laguerre, 392
  Hermite, 274
  Laguerre, 392
  Legendre, 337, 380
Position representation, 142–143
Positron, 25, 764
Positronium, 26
Postulates of quantum mechanics, 187–190
Potential
  barrier and well, 250–258
  central, 374–377
  centrifugal or effective, 376
  Coulomb, 386
  delta, 285
  double-delta, 288
  finite square well, 262–266
  harmonic oscillator, 266–277
  infinite square well, 258–262
  spherical square well, 380
  step, 246–250
  Yukawa, 693
Power radiated, 649
Probabilistic interpretation, 40
Probability
  current, 205, 755
  density, 40, 205, 723, 755
Proca equation, 930
Propagation of wave packets, 55–66
Pseudo-scalar, 471
Pseudo-scalars, 805
Pseudo-vector, 471
Pseudo-vectors, 805

QED Lagrangian density, 921
Quadratic Stark effect, 543
Quadrupole interaction, 600, 601
Quantization of action, 48
Quantization of electromagnetic field, 641–644
Quantization rule
  Bohr, 41
  Bohr–Sommerfeld, 573
  Planck, 16
  Wilson–Sommerfeld, 48
Quantization Rules, 48–50
Quantum Electrodynamics, 920
Quantum electrodynamics, 919
Quantum Field Theory, 919
Quantum field theory, 819, 897
Quantum number, 260, 262
Quantum theory of radiation, 641–644
Quarks, 512

Radial equation for
  a central potential, 376
  a free particle, 378
  a hydrogen atom, 388
  an isotropic oscillator, 382
Raising operator, 269
Ramsauer–Townsend effect, 254
Rapidity, 963
Rayleigh–Jeans formula, 16
Rayleigh–Ritz method, 558
Reduced mass, 676
Reduced matrix element, 477
Reducible tensors, 474
Reflection coefficient, 248, 867
Relativistic corrections, 551
Relativistic Notation, 961
Residue theorem, 579
Rigid rotor, 345
Rodrigues formula, 337
Rotation group, 440
Rotation operator, 440
Rotations
  and spherical harmonics, 445–448
  classical, 433–437
  Euler, 442–443
  finite, 439–440
  in quantum mechanics, 437–448
  infinitesimal, 438–439
Rutherford model, 41
Rutherford scattering formula, 684
Rydberg constant, 43, 390

Scalar field, 911
Scalar operator, 471
Scalar particle, 725, 911
Scalar product, 92, 96–98, 143, 144
Scattering
  amplitude, 675–682
  cross section, 671–675
  of identical particles, 691–693
  Rutherford, 684
Schrödinger equation
  numerical solution of, 277
  time-dependent, 189, 202, 203, 624
  time-independent, 202, 203, 241
Schrödinger picture, 624
Schrödinger–Pauli equation, 407
Schrödinger-Pauli equation, 796
Schwarz inequality, 98
Screening effect, 522
Second quantization, 644
Secular equation, 138
Selection rules
  Clebsch–Gordan coefficients, 453
  dipole transitions, 490, 646–647
Self-conjugate particles, 818
Semiclassical approximation, 566
Separation of variables, 367–369, 374, 376
Shell structure, 518
Simultaneous eigenstates, 199
Simultaneous measurements, 199, 207
Simultaneous observables, 199
Singlet state, 457, 470, 517, 530, 602, 692
Slater determinant, 515
Space
  dimension of a, 94
  Euclidean, 94
  Hilbert, 92
  linear vector, 91
  phase, 49, 574
Spacelike, 713, 967
Special $SO(3)$ group, 437, 445
Spectroscopic notation, 520
Spherical Hankel functions, 381
Spherical harmonics, 339–343
Spin
  angular momentum, 325–333
  experimental evidence, 325
  general theory, 329–330
  magnetic dipole moment, 327
  Singlet state, 457
  Triplet state, 457
spin magnetic moment, 407, 796
Spin–orbit coupling, 548
Spin–orbit functions, 463
Spin–orbit interaction, 549
Spin-statistics theorem, 7
Spinors
  Pauli, 331
Spontaneous emission, 645, 647–650
Square-integrable functions, 96
Standard model, 920
Standard representation, 747
Stark effect, 543, 546
  linear, 543
  quadratic, 543
States
  orthogonal, 99
  orthonormal, 99
Stationary states, 41, 202–203, 241
Stefan–Boltzmann constant, 14
Stefan–Boltzmann law, 14
Stern–Gerlach experiment, 6–329
Stimulated emission, 633, 645
Sudden approximation, 636–638
Superposition principle, 37, 191
Symmetric state, 506, 510–512
Symmetrization postulate, 511

Tensor operators, 471–479
The 4-Vector Notation, 964
Thomas precession, 549
Thomson experiment, 29
Time evolution operator, 201
Time reveral operator, 117
Time reversal, 114
Time–energy uncertainty relation, 39
Time-dependent perturbation theory, 627–634
Time-independent perturbation theory, 538–558
Timelike, 713, 967
Transformation theory, 9
Transition rate, 631, 641
Transmission coefficient, 248, 867
Triangle inequality, 99

Triplet state, 457, 471, 517, 530, 602, 692, 693
Tunneling, 254, 256–258, 579–582

Uhlenbeck, George, 326
Ultraviolet catastrophe, 16
Uncertainty principle, 11
Uncertainty relations, 38, 39, 111
Unit matrix, 126
Unitary matrix, 128
Unitary operators, 113
Unitary transformations, 120–123
  finite, 123, 208–209
  infinitesimal, 122, 207–208
  properties, 121

Variational method, 558–566
Vector operator, 472
Vectors
  linearly dependent, 94
  linearly independent, 93
Velocity
  group, 56, 57
  phase, 56
Virial theorem, 43, 277, 399
von Neumann, 12

Wave function
  antisymmetric, 512
  Continuity condition, 249
  of many-particle systems, 515
  of three-particle systems, 514
  of two-particle systems, 513
  symmetric, 512
Wave mechanics, 151
Wave packet, 50–66
  distorted, 59–64
  Gaussian, 52, 55
  group velocity, 57
  localized, 51
  minimum uncertainty, 55
  motion of a, 55
  Schrödinger equation and, 203
  spreading of a, 60
  time evolution, 58–59
  undistorted, 55
Wave vector, 12
Wave–particle duality, 35–36
Wavelength, de Broglie, 29
Weyl equations, 784
Weyl fermions, 786, 787
Weyl representation, 747
Weyl spinors, 785
Wien's displacement law, 18
Wien's formula, 14
Wigner *D*-matrix, 443
Wigner formula, 444
Wigner functions, 443
Wigner Theorem, 116
Wigner theorem, 811
Wigner–Eckart theorem, 477
Wilson–Sommerfeld quantization rule, 48
WKB method, 566–582
  applied to tunneling, 579
  connection formulas, 572, 573
  for bound states, 569–578
  for central potentials, 574
  for Coulomb potentials, 578
  quantization condition, 573
Work function, 20
World line, 965

Yukawa potential, 693

Zeeman Effect, 4
Zeeman effect
  anomalous, 555–558
  normal, 401–403
  strong-field, 556
  weak-field, 556–558
Zero-point energy, 260, 271, 574
Zitterbewegung effect, 798

Physical Constants

Quantity	Symbol, equation	Value
Speed of light	c	$2.997\,9 \times 10^{8}\,\mathrm{m\,s^{-1}}$
Electron charge	e	$1.602 \times 10^{-19}\,\mathrm{C}$
Planck constant	h	$6.626 \times 10^{-34}\,\mathrm{J\,s} = 4.136 \times 10^{-15}\,\mathrm{eV\,s}$
Reduced Planck constant	$\hbar = h/2\pi$	$1.055 \times 10^{-34}\,\mathrm{J\,s} = 6.582 \times 10^{-16}\,\mathrm{eV\,s}$
Conversion constant	$\hbar c$	$197.327\,\mathrm{MeV\,fm} = 197.327\,\mathrm{eV\,nm}$
Electron mass	m_e	$9.109 \times 10^{-31}\,\mathrm{kg} = 0.511\,\mathrm{MeV/c^2}$
Proton mass	m_p	$1.673 \times 10^{-27}\,\mathrm{kg} = 938.272\,\mathrm{MeV/c^2}$
Neutron mass	m_n	$1.675 \times 10^{-27}\,\mathrm{kg} = 939.566\,\mathrm{MeV/c^2}$
Fine structure constant	$\alpha = e^2/(4\pi\varepsilon_0\hbar c)$	$1/137.036$
Classical electron radius	$r_e = e^2/(4\pi\varepsilon_0 m_e c^2)$	$2.818 \times 10^{-15}\,\mathrm{m}$
Electron Compton wavelength	$\lambda_e = h/m_e c = r_e/\alpha$	$2.426 \times 10^{-12}\,\mathrm{m}$
Proton Compton wavelength	$\lambda_p = h/m_p c$	$1.321 \times 10^{-15}\,\mathrm{m}$
Bohr radius	$a_0 = r_e/\alpha^2$	$0.529 \times 10^{-10}\,\mathrm{m}$
Rydberg energy	$\mathcal{R} = m_e c^2 \alpha^2/2$	$13.606\,\mathrm{eV}$
Electron speed in the first Bohr orbit	$v_1 = \alpha c$	$2.1876913 \times 10^{6}\,\mathrm{m/s}$
Bohr magneton	$\mu_B = e\hbar/2m_e$	$5.788 \times 10^{-11}\,\mathrm{MeV\,T^{-1}}$
Nuclear magneton	$\mu_N = e\hbar/2m_p$	$3.152 \times 10^{-14}\,\mathrm{MeV\,T^{-1}}$
Avogadro number	N_A	$6.022 \times 10^{23}\,\mathrm{mol^{-1}}$
Boltzmann constant	k	$1.381 \times 10^{-23}\,\mathrm{J\,K^{-1}} = 8.617 \times 10^{-5}\,\mathrm{eV\,K^{-1}}$
Stefan–Boltzmann constant	$2\pi^5 k^4/(15h^3c^2)$	$5.67 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$
Gas constant	$R = N_A k$	$8.31\,\mathrm{J\,mol^{-1}\,K^{-1}}$
Gravitational constant	G	$6.673 \times 10^{-11}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$
Permeability of free space	μ_0	$4\pi \times 10^{-7}\,\mathrm{T\,m\,A^{-1}}$
Permittivity of free space	$\varepsilon_0 = 1/\mu_0 c^2$	$8.854 \times 10^{-12}\,\mathrm{C^2\,N^{-1}\,m^{-2}}$
Useful quantity	$e^2/(4\pi\varepsilon_0)$	$1.440\,\mathrm{MeV\,fm}$

Conversion of units

$1\,\mathrm{fm} = 10^{-15}\,\mathrm{m}$, $1\,\mathrm{barn} = 10^{-28}\,\mathrm{m^2} = 100\,\mathrm{fm^2}$, $1\,\mathrm{G} = 10^{-4}\,\mathrm{T}$

$1\,\mathrm{atmosphere} = 101\,325\,\mathrm{Pa}$ Thermal energy at $T = 300\,\mathrm{K}$: $kT = [38.682]^{-1}\,\mathrm{eV}$

$0\,^{\circ}\mathrm{C} = 273.15\,\mathrm{K}$, $1\,\mathrm{eV} = 1.602 \times 10^{-19}\,\mathrm{J}$, $1\,\mathrm{eV/c^2} = 1.783 \times 10^{-36}\,\mathrm{kg}$

Essential Relations

Planck's Distribution: $u(\nu, T) = (8\pi\nu^2/c^3)\left(h\nu/\left(e^{h\nu/kT} - 1\right)\right)$, **de Broglie Relation:** $\lambda = h/p$

Compton wavelength shift: $\Delta\lambda = h(1 - \cos\theta)/m_e c$, **Fine structure constant:** $\alpha = \left(\frac{e^2}{4\pi\epsilon_0}\right)\frac{1}{\hbar c}$

Bohr model: $a_0 = \left(\frac{4\pi\epsilon_0}{e^2}\right)\frac{\hbar^2}{m_e}$, $r_n = n^2 a_0$, $v_n = \frac{a}{n}c$, $E_n = -\left(\frac{e^2}{4\pi\epsilon_0}\right)\frac{1}{2n^2 a_0} \equiv -\frac{1}{2n^2}\alpha^2 m_e c^2$

General relations:

$$\begin{aligned} e^{\hat{A}}e^{\hat{B}} &= e^{\hat{A}+\hat{B}}e^{[\hat{A},\,\hat{B}]/2}, \\ e^{\hat{A}}\hat{B}e^{-\hat{A}} &= \hat{B} + [\hat{A},\,\hat{B}] + \frac{1}{2!}[\hat{A},\,[\hat{A},\,\hat{B}]] + \frac{1}{3!}[\hat{A},\,[\hat{A},\,[\hat{A},\,\hat{B}]]] + \cdots \end{aligned}$$

Generalized uncertainty principle: $\Delta A \Delta B \geq \frac{1}{2}\left|\langle[\hat{A},\,\hat{B}]\rangle\right|$, where $\Delta A = \sqrt{\langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2}$

Canonical commutator: $[\hat{x},\,\hat{p}] = i\hbar$

Heisenberg uncertainty principle: $\Delta x \Delta p \geq \hbar/2$, $\Delta E \Delta t \geq \hbar/2$

Measurement probability: $\hat{A}|\psi_n\rangle = a_n|\psi_n\rangle$, $P_n(a_n) = \dfrac{|\langle\psi_n|\psi\rangle|^2}{\langle\psi|\psi\rangle}$

Expectation value: $\langle\hat{A}\rangle = \dfrac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} = \sum_n a_n P_n(a_n)$

Time evolution of expectation values: $\dfrac{d}{dt}\langle\hat{A}\rangle = \dfrac{1}{i\hbar}\langle[\hat{A},\hat{H}]\rangle + \langle\dfrac{\partial\hat{A}}{\partial t}\rangle$

Commutators and Poisson brackets: $\dfrac{1}{i\hbar}[\hat{A},\hat{B}] \longrightarrow \{A,B\}_{classical}$

Time-dependent Schrödinger equation: $i\hbar\dfrac{\partial|\Psi(t)\rangle}{\partial t} = \hat{H}|\Psi(t)\rangle$

Probability density: $\rho(\vec{r},t) = \Psi^*(\vec{r},t)\Psi(\vec{r},t)$

Probability current density: $\vec{j}(\vec{r},t) = \dfrac{i\hbar}{2m}(\Psi\vec{\nabla}\Psi^* - \Psi^*\vec{\nabla}\Psi)$

Conservation of probability: $\dfrac{\partial\rho(\vec{r},t)}{\partial t} + \vec{\nabla}\cdot\vec{j} = 0$

Angular momentum:

$$\begin{aligned} [\hat{J}_x,\hat{J}_y] &= i\hbar\hat{J}_z, \qquad [\hat{J}_y,\hat{J}_z] = i\hbar\hat{J}_x, \qquad [\hat{J}_z,\hat{J}_x] = i\hbar\hat{J}_y \\ \hat{J}^2\,|\,j,\,m\rangle &= \hbar^2 j(j+1)\,|\,j,\,m\rangle, \qquad \hat{J}_z\,|\,j,\,m\rangle = \hbar m\,|\,j,\,m\rangle \\ \hat{J}_\pm\,|\,j,\,m\rangle &= \hbar\sqrt{j(j+1) - m(m\pm1)}\;|\,j,\,m\pm1\rangle \\ \langle j,\,m\,|\,\hat{J}_x^2\,|\,j,\,m\rangle &= \langle j,\,m\,|\,\hat{J}_y^2\,|\,j,\,m\rangle = \frac{\hbar^2}{2}\left[j(j+1) - m^2\right] \end{aligned}$$

For $j = \frac{1}{2}$: $J_k = \frac{\hbar}{2}\sigma_k$ (with $k = x, y, z$), where σ_x, σ_y, and σ_z are the Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

For $j = 1$: the matrices of J_x, J_y, and J_z are

$$\hat{J}_x = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \hat{J}_y = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \hat{J}_z = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Time-independent potentials: $|\Psi(t)\rangle = |\psi(x)\rangle \exp(-iEt/\hbar)$

Time-independent Schrödinger equation: $-\frac{\hbar^2}{2m}\nabla^2\psi(x) + V(x)\psi(x) = E\psi(x)$

Infinite square well: $E_n = -\frac{\hbar^2\pi^2}{2ma^2}n^2, \qquad \psi_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right) \qquad (n = 1, 2, 3, \ldots)$

Harmonic oscillator:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{X}^2 = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right), \qquad E_n = \hbar\omega\left(n + \frac{1}{2}\right)$$

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \qquad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \qquad [\hat{a}, \hat{a}^\dagger] = 1$$

$$\langle n \mid \hat{X}^2 \mid n\rangle = \frac{\hbar}{2m\omega}(2n+1), \qquad \langle n \mid \hat{P}^2 \mid n\rangle = \frac{m\hbar\omega}{2}(2n+1)$$

Hydrogen atom: radial equation and averages: $-\frac{\hbar^2}{2\mu}\frac{d^2U(r)}{dr^2} + \left[\frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{e^2}{r}\right]U(r) = EU(r)$

$$\langle nl|\, r\,|nl\rangle = \frac{1}{2}[3n^2 - l(l+1)]a_0, \qquad \langle nl|\, r^2 \mid nl\rangle = \frac{1}{2}n^2[5n^2 + 1 - 3l\,(l+1)]a_0^2$$

$$\langle nl\,|r^{-1}|\,nl\rangle = \frac{1}{n^2\,a_0}, \qquad \langle nl|r^{-2}|nl\rangle = \frac{2}{n^3(2l\,+1)a_0^2}$$

Time-independent perturbation theory:

$$(\hat{H}_0 + \hat{H}_p) \mid \psi_n\rangle = E_n \mid \psi_n\rangle, \qquad \hat{H}_0 \mid \phi_n\rangle = E_n^{(0)} \mid \phi_n\rangle \qquad (\hat{H}_p \ll \hat{H}_0)$$

$$E_n = E_0 + \langle\phi_n \mid \hat{H}_p \mid \phi_n\rangle + \sum_{m\neq n}\frac{\left|\langle\phi_m|\hat{H}_p \mid \phi_n\rangle\right|^2}{E_n^{(0)} - E_m^{(0)}} + \cdots, \quad |\psi_n\rangle = |\phi_n\rangle + \sum_{m\neq n}\frac{\langle\phi_m \mid \hat{H}_p \mid \phi_n\rangle}{E_n^{(0)} - E_m^{(0)}}\mid\phi_m\rangle + \cdots$$

Quantization condition: $\oint p(x, E_n)\,dx = 2\int_{x_1}^{x_2}\sqrt{2m(E_n - V(x))}\,dx = (n + \frac{1}{2})h$

Time-dependent potentials: Heisenberg and interaction pictures:

$$|\psi(t)\rangle_H = e^{it\hat{H}/\hbar} \mid \psi(t)\rangle, \quad \hat{A}_H(t) = e^{it\hat{H}/\hbar}\hat{A}e^{-it\hat{H}/\hbar}, \quad \frac{d\hat{A}_H}{dt} = \frac{1}{i\hbar}[\hat{A}_H, \hat{H}]$$

$$|\psi(t)\rangle_I = e^{i\hat{H}_0t/\hbar} \mid \psi(t)\rangle, \quad \hat{V}_I(t) = e^{i\hat{H}_0t/\hbar}\hat{V}e^{-i\hat{H}_0t/\hbar}, \quad i\hbar\frac{d \mid \psi(t)\rangle_I}{dt} = \hat{V}_I(t) \mid \psi(t)\rangle_I$$

Time-dependent perturbation theory: $P_{if}(t) = \left|-\frac{i}{\hbar}\int_0^t \langle\psi_f \mid \hat{V}(t') \mid \psi_i\rangle e^{i\omega_{fi}t'}\,dt'\right|^2$

Intensity of radiation emitted: $I_{i\to f} = \hbar\omega W_{i\to f}^{emi} = \frac{4}{3}\frac{\omega^4}{c^3}|\vec{d}_{fi}|^2 = \frac{4}{3}\frac{\omega^4 e^2}{c^3}\left|\langle\psi_f \mid \vec{r}|\psi_i\rangle\right|^2$

Scattering:

Differential cross section (Born approximation): $\frac{d\sigma}{d\Omega} = \left|f(\theta,\varphi)\right|^2 = \frac{\mu^2}{4\pi^2\hbar^4}\left|\int e^{i\vec{q}\cdot\vec{r}'}V(\vec{r}')\,d^3r'\right|^2$

Partial wave analysis: $f(\theta) = \sum_{l=0}^{\infty} f_l(\theta) = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)e^{i\delta_l}\sin\delta_l P_l(\cos\theta)$

Relativistic Quantum Mechanics:

Klein–Gordon Equation: $\left(\partial_\mu\partial^\mu + \frac{m^2c^2}{\hbar^2}\right)\psi(x^\mu) = 0$, **Dirac Equation:** $\left(i\hbar\gamma^\mu\partial_\mu - mc\right)\psi(x^\mu) = 0$

An innovative approach to quantum mechanics that seamlessly combines textbook and problem-solving book into one

Quantum Mechanics: Concepts and Applications provides an in-depth treatment of this fundamental theory, combining detailed formalism with straightforward practice. Thoroughly integrating close to seven hundred examples, solved problems, and exercises into a well-structured and comprehensive work, this textbook offers instructors a pedagogically sound teaching tool, students a clear, balanced and modern approach to the subject, and researchers a quick practical guide. The extensive list of fully solved examples and problems have been carefully designed to guide and enable users of the book to become proficient practitioners of quantum mechanics.

The text begins with a thorough description of the origins of quantum physics before discussing the mathematical tools required in the field and the postulates upon which it is founded. *Quantum Mechanics: Concepts and Applications* is broad in scope, covering such aspects as one-dimensional and three-dimensional potentials, angular momentum, rotations and addition of angular momenta, identical particles, time-independent and -dependent approximation methods, scattering theory, relativistic quantum mechanics, and classical field theory among others. Each of these diverse areas are enhanced with a rich collection of illustrative examples and fully-solved problems to ensure complete understanding of this complex topic.

Readers of the third edition of *Quantum Mechanics: Concepts and Applications* will also find:

- Two new chapters — one dealing with relativistic quantum mechanics and the other with the Lagrangian derivations of the Klein-Gordon and Dirac equations — and three new appendices to support them
- About 90 solved examples integrated throughout the text that are intended to illustrate individual concepts within a broader topic
- About 200 fully-solved, multi-step problems at the end of each chapter that integrate multiple concepts introduced throughout the chapter
- More than 400 unsolved exercises that may be used to practice the ideas presented

A Solutions Manual is available only to those instructors adopting the book, on request, offering detailed solutions to all exercises.

Quantum Mechanics: Concepts and Applications is a comprehensive textbook which is most useful to senior undergraduate and first-year graduate students seeking mastery of the field, as well as to researchers in need of a quick, practical reference for the various techniques necessary for optimal performance in the subject.

Nouredine Zettili, PhD, is Professor of Physics at Jacksonville State University, Alabama, USA. He received his MS and PhD in theoretical nuclear physics from MIT. His research interests include nuclear theory, the many-body problem, quantum mechanics, and mathematical physics.

Cover Design: Wiley
Cover Image: © ALFRED PASIEKA/Getty Images

www.wiley.com

WILEY

ISBN 978-1-118-30789-2
90000
9 781118 307892